

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2006, 17:51:20 ; Search time 10461 Seconds
(without alignments)
17149.214 Million cell updates/sec

Title: US-10-681-086-1
Perfect score: 3156
Sequence: 1 ggatccacgaggtacgac.....atcggtatgatgcgaattc 3156

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result/No.	Score	Query Match	Length	DB	ID	Description
1	3156	100.0	3156	6	BD175786	BD175786 Process f
2	3156	100.0	3156	6	AR435692	AR435692 Sequence
3	3156	100.0	3156	6	AX394258	AX394258 Sequence
4	1612.6	51.1	8478	6	AR173058	AR173058 Sequence
5	1612.6	51.1	8478	6	AR630545	AR630545 Sequence
6	1612.6	51.1	10153	1	BSU51868	US1868 Bacillus su
7	1612.6	51.1	199922	1	BSUB0016	Z99119 Bacillus su
8	1612.6	51.1	220060	1	AF008220	AF008220 Bacillus
9	1431	45.3	6739	1	AB088066	AB088066 Bacillus
10	1219	38.6	3560	6	AX825002	AX825002 Sequence
11	1200.4	38.0	7350	6	CQ894464	CQ894464 Sequence
12	1192.8	37.8	8152	6	CS115035	CS115035 Sequence
13	1192.8	37.8	8644	6	CS115022	CS115022 Sequence
14	1192.8	37.8	13129	6	CS091382	CS091382 Sequence
15	867	27.5	4335	6	CS119875	CS119875 Sequence
16	867	27.5	4545	1	PB0KANRCG	M37273 Plasmid pub
17	867	27.5	4548	1	PB0110CG	M19465 Plasmid pub
18	867	27.5	4780	11	ACVPBD64	X51450 Artificial

19	867	27.5	5115	1	AB037420	AB037420 Staphyloc
20	867	27.5	5142	6	AX069289	AX069289 Sequence
21	867	27.5	5793	6	AR569128	AR569128 Sequence
22	867	27.5	5943	6	AR569127	AR569127 Sequence
23	867	27.5	6561	6	AX601433	AX601433 Sequence
24	867	27.5	7311	6	CS091381	CS091381 Sequence
25	867	27.5	7336	6	A08030	A08030 Synthetic n
26	867	27.5	7336	6	A13198	A13198 complete pl
27	867	27.5	7336	6	A20502	A20502 Hind III fr
28	867	27.5	7336	6	AR033984	AR033984 Sequence
29	867	27.5	7336	6	AR123993	AR123993 Sequence
30	867	27.5	7336	6	AR154323	AR154323 Sequence
31	867	27.5	7336	6	AR408644	AR408644 Sequence
32	867	27.5	8119	6	AR002333	AR002333 Sequence
33	867	27.5	8119	6	AR018032	AR018032 Sequence
34	867	27.5	8119	6	AR055698	AR055698 Sequence
35	867	27.5	9144	6	CQ815355	CQ815355 Sequence
36	867	27.5	46445	1	AF051917	AF051917 Staphyloc
37	867	27.5	58237	1	D86934	D86934 Staphylococ
38	867	27.5	110000	1	BA000017_00	BA000017 Staphyloc
39	867	27.5	110000	1	BA000018_00	BA000018 Staphyloc
40	867	27.5	110000	1	EX571856_00	EX571856 Staphyloc
41	865.4	27.4	1200	1	PTBKAN	K02551 Plasmid pTB
42	865.4	27.4	2262	1	BSPRBHJK	X03409 Thermophili
43	865.4	27.4	4525	1	BSREPB	X15670 Bacillus pl
44	865.4	27.4	5767	6	AX035965	AX035965 Sequence
45	865.4	27.4	6447	11	AF288421	AF288421 Synthetic

ALIGNMENTS

RESULT 1
BD175786
LOCUS BD175786 3156 bp DNA linear PAT 18-MAR-2003
DEFINITION Process for producing target fermentation product.
ACCESSION BD175786
VERSION BD175786.1 GI:29121488
KEYWORDS JP 2002253273-A/1.
SOURCE Bacillus subtilis
ORGANISM Bacillus subtilis
REFERENCE 1 (bases 1 to 3156)
AUTHORS Hohmann,H.P., Mouncey,N.J., Schlieker,H.W. and Stebbins,J.W.
TITLE Process for producing target fermentation product
JOURNAL Patent: JP 2002253273-A 1 10-SEP-2002;
ROCHE VITAMINS AG
COMMENT OS Bacillus subtilis
PN JP 2002253273-A/1
PD 10-SEP-2002
PF 08-AUG-2001 JP 2001240984
PR 08-AUG-2000 US 09/633927
PI HANS PETER HOHMANN,NIGEL JOHN MOUNCEY,HEINRICH WINFRIED PI
SCHLIEKER,
PI JEFFREY W STEBBINS
PC C12N15/09,C12N1/21,C12P25/00//C12N1/21,C12R1:125), (C12P25/00,
PC C12R1:125),
PC C12N15/00,
CC Process for producing target fermentation product PH Key
FT Location/Qualifiers
FT source 1. 3156
FEATURES
source Location/Qualifiers
1. 3156 /organism="Bacillus subtilis"
/mol_type="genomic DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 100.0%; Score 3156; DB 6; Length 3156;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCAGAGGTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTAGACAGA 60
DB 1 GGATCCAGAGGTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTAGACAGA 60
QY 61 ATGAAAGAGCGCGGTACATCGTAAACCTGCGGTCAATGGAATGGAGCGCCGGTCCAGAG 120
DB 61 ATGAAAGAGCGCGGTACATCGTAAACCTGCGGTCAATGGAATGGAGCGCCGGTCCAGAG 120
QY 121 AGGAATATTGATGGGAAATCAACCGCTCTGGTCTCAACCAATATTATAGGCGTCGA 180
DB 121 AGGAATATTGATGGGAAATCAACCGCTCTGGTCTCAACCAATATTATAGGCGTCGA 180
QY 181 AGCGATAGACGTTTGCATCGATCGAGCCCAACAGCAATTCGACGAATTTGGGACAGGAGC 240
DB 181 AGCGATAGACGTTTGCATCGATCGAGCCCAACAGCAATTTGGGACAGGAGC 240
QY 241 AGCGGTTCACTGTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATT 300
DB 241 AGCGGTTCACTGTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATT 300
QY 301 GCCAGCTTTAACTGACAGAGCGCCCTGCTGTTTCGAGCGGTACTTGGCCCAATGTC 360
DB 301 GCCAGCTTTAACTGACAGAGCGCCCTGCTGTTTCGAGCGGTACTTGGCCCAATGTC 360
QY 361 GGTGTCTTTTCATCTTCCAGAAAGGAAGATGTCAATTTAAAGTGACGAGCTCAATCAT 420
DB 361 GGTGTCTTTTCATCTTCCAGAAAGGAAGATGTCAATTTAAAGTGACGAGCTCAATCAT 420
QY 421 GCAAGTATGATCGAGCGGTGCGGACTTCTTAAGGCTGATACAGTGTGTTATCGGCATATT 480
DB 421 GCAAGTATGATCGAGCGGTGCGGACTTCTTAAGGCTGATACAGTGTGTTATCGGCATATT 480
QY 481 CATATGAATGATCTTGAAACCAAGCTGAATGAAACACAGCGTTATCAGCGCGGTTTATC 540
DB 481 CATATGAATGATCTTGAAACCAAGCTGAATGAAACACAGCGTTATCAGCGCGGTTTATC 540
QY 541 GTAAACAGCGGATATTCAAGATGATGGCACAAATCGCCCTCTTGATCAGATCATCTCA 600
DB 541 GTAAACAGCGGATATTCAAGATGATGGCACAAATCGCCCTCTTGATCAGATCATCTCA 600
QY 601 CTTGCGAAACGCTATCATGCTTCTGTTGCTGTTGATGATGCCACGCAACAGGAGTTTG 660
DB 601 CTTGCGAAACGCTATCATGCTTCTGTTGCTGTTGATGATGCCACGCAACAGGAGTTTG 660
QY 661 GCGGATTCGGGACAGGAAACAGTGAATACCTTGGTGTGTTGCTCCGACATGTTATCGGC 720
DB 661 GCGGATTCGGGACAGGAAACAGTGAATACCTTGGTGTGTTGCTCCGACATGTTATCGGC 720
QY 721 ACCTTAAGCAAAGCTGTTGGCGGAAAGAGGTTTTCGGCAGGATCAGCGGCTCTTCATC 780
DB 721 ACCTTAAGCAAAGCTGTTGGCGGAAAGAGGTTTTCGGCAGGATCAGCGGCTCTTCATC 780
QY 781 GACTTTTTCGTGAACCATGCGAGAACATTTATCTTTTCAAAACCGCTATTCGCCAGCCAGC 840
DB 781 GACTTTTTCGTGAACCATGCGAGAACATTTATCTTTTCAAAACCGCTATTCGCCAGCCAGC 840
QY 841 TGTGCGGCTGTCACGAGGCTTTCAACATATTGAAAGCAAGGAAACGACAGCTT 900
DB 841 TGTGCGGCTGTCACGAGGCTTTCAACATATTGAAAGCAAGGAAACGACAGCTT 900
QY 901 TTATTTTCTTATATCAGCATGATCAGAACAGTCTGAAGAAATATGGGTATATGGTGA 960
DB 901 TTATTTTCTTATATCAGCATGATCAGAACAGTCTGAAGAAATATGGGTATATGGTGA 960
QY 961 GGAGATCACACACCGATTTATTCCTGTAGTCAATGGCGATGCCATAAAGCGTCTTATTT 1020
DB 961 GGAGATCACACACCGATTTATTCCTGTAGTCAATGGCGATGCCATAAAGCGTCTTATTT 1020
QY 1021 GCTGAATACTGACGGGCAAGGAATTTATGCTCTCGCATTCGGCCGCAACCGTTGCG 1080
DB 1021 GCTGAATACTGACGGGCAAGGAATTTATGCTCTCGCATTCGGCCGCAACCGTTGCG 1080
QY 1081 CCGGGTGAAGCCGGAATTCGAAGCTTGGGACAGAGTTCGAGATCAGGGAATGAGTTTATA 1140

DB 1081 CCGGGTGAAGCCGGATTCGAAGCTTGGGACAGAGTTCGAGATCAGGGAATGAGTTTATA 1140
QY 1141 AAATAAAAAAGCACCTGAAAAAGGTGCTTTTTTTGATGGTTTGAACCTGTTCTTCTT 1200
DB 1141 AAATAAAAAAGCACCTGAAAAAGGTGCTTTTTTTGATGGTTTGAACCTGTTCTTCTT 1200
QY 1201 ATCTTTGATACATATAGAAATAACGTCATTTTTTTATTTTTATTTTGTCTGAAAGGTGCG 1260
DB 1201 ATCTTTGATACATATAGAAATAACGTCATTTTTTTATTTTTATTTTTAGTTGCTGAAAGGTGCG 1260
QY 1261 TTGAAGTGTGGTATGATGTTTAAAGTATTGAAAAACCCCTTAAAAATTTGGTTGCAAG 1320
DB 1261 TTGAAGTGTGGTATGATGTTTAAAGTATTGAAAAACCCCTTAAAAATTTGGTTGCAAG 1320
QY 1321 AAAAACCCCATCTGTTTAAAGTTATAAGTGACTAAACAAATAAATAAATAGATGGGGTTT 1380
DB 1321 AAAAACCCCATCTGTTTAAAGTTATAAGTGACTAAACAAATAAATAAATAGATGGGGTTT 1380
QY 1381 CTTTTTAATATTATGTGCTCTAAATAGTAGCATTTTCCAGATGAAAAATCAAGGTTTTAG 1440
DB 1381 CTTTTTAATATTATGTGCTCTAAATAGTAGCATTTTATCAGATGAAAAATCAAGGTTTTAG 1440
QY 1441 TGGAACAAGAAAAAGTGGAAAAAGTGAGACCATGTGCTTAGGAAGACGAGTTATTAATAG 1500
DB 1441 TGGAACAAGAAAAAGTGGAAAAAGTGAGACCATGTGCTTAGGAAGACGAGTTATTAATAG 1500
QY 1501 CTGATTAAGAAACGGTGTCTCCAAATATTCTTATTTAGAAAAACCAATCTAAAAATTATCT 1560
DB 1501 CTGAAATGAAGACGGTGTCTCTCCAAATATTCTTATTTAGAAAAACCAATCTAAAAATTATCT 1560
QY 1561 GAAAAAGGAATGAGATGAGTGAATGGACCAATTAATATGACTAGAGAAAGAAATGAAG 1620
DB 1561 GAAAAAGGAATGAGAAATGAGATGAGTGAATGGACCAATTAATATGACTAGAGAAAGAAATGAAG 1620
QY 1621 ATTGTTTCATGAAATTAAGGAAACGAATATTGGATAAATATGGGGATGATGTTTAAGGCTATT 1680
DB 1621 ATTGTTTCATGAAATTAAGGAAACGAATATTGGATAAATATGGGGATGATGTTTAAGGCTATT 1680
QY 1681 GGTGTTTTATGGCTCTCTTGGTCTGTCAGACTGATGGGCCCTTATTCGGATATTGAGATGATG 1740
DB 1681 GGTGTTTTATGGCTCTCTTGGTCTGTCAGACTGATGGGCCCTTATTCGGATATTGAGATGATG 1740
QY 1741 TGTGTCATGTCACACAGAGGAGCAGAGTTTCAGCCATGAAATGGACCAACCGGTGAGTGAAG 1800
DB 1741 TGTGTCATGTCACACAGAGGAGCAGAGTTTCAGCCATGAAATGGACCAACCGGTGAGTGAAG 1800
QY 1801 GTGGAAGTGAATTTTGAATAGCAGAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCA 1860
DB 1801 GTGGAAGTGAATTTTGAATAGCAGAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCA 1860
QY 1861 GATTGGCCGCTTACACATGTCATTTTTTCTCTATTTTGGGATTTATGATTCAAGTGGA 1920
DB 1861 GATTGGCCGCTTACACATGTCATTTTTTCTCTATTTTGGGATTTATGATTCAAGTGGA 1920
QY 1921 TACTTAGAGAAAGTGTATCAAACTGCTTAAATCGGTAGAAGCCCAAGCTTCCAGATGCG 1980
DB 1921 TACTTAGAGAAAGTGTATCAAACTGCTTAAATCGGTAGAAGCCCAAGCTTCCAGATGCG 1980
QY 1981 ATTTGTGCCCTTATCGTAGAAGAGCTGTTTTGAATATGACGCAAAATGGCGTAATTTTCGT 2040
DB 1981 ATTTGTGCCCTTATCGTAGAAGAGCTGTTTTGAATATGACGCAAAATGGCGTAATTTTCGT 2040
QY 2041 GTGCAAGGACCGACACATTTCTACCATCTTGTGATCAGGTAGCAATGGCAGGTGCC 2100
DB 2041 GTGCAAGGACCGACACATTTCTACCATCTTGTGATCAGGTAGCAATGGCAGGTGCC 2100
QY 2101 ATGTGATTCGCTGTCATCATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
DB 2101 ATGTGATTCGCTGTCATCATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
QY 2161 GCAGTTAAGCAATCAGATCTTCCCTCAGGTTTATGACCATCTGTGCCAGTTTCGTAAATGCT 2220

Db	2161	GCAGTTAAGCAATCAGATCTTCTCTTTCAGGTTATGACCAATCTGTGCCAGTTTCGTAATGTCT	2222
Qy	2221	GGTCAACTTTCCGACTCTGAGAACTCTTTCGGAATCGCTAGAGAAATTTCTCGAAATGGGATT	2280
Db	2221	GGTCAACTTTCCGACTCTGAGAACTCTTTCGGAATCGCTAGAGAAATTTCTCGAAATGGGATT	2280
Qy	2281	CAGAGTGGACAGAAACGACACAGGATATATAGTGGATGTGTCAAAAACGATACCAATTTTGA	2340
Db	2281	CAGAGTGGACAGAAACGACACAGGATATATAGTGGATGTGTCAAAAACGATACCAATTTTGA	2340
Qy	2341	ATTCCGAAGCGCCGATTTGAGTCTTACCGGATGGTGAATAAGGAAAACGCTCTTTGAAGGCG	2400
Db	2341	ATTCCGAAGCGCCGATTTGAGTCTTACCGGATGGTGAATAAGGAAAACGCTCTTTGAAGGCG	2400
Qy	2401	CGAAGCGGGCGCACGATCTGAATATCGGCACATATTTGATCGTGCGCAAGCGGACAGGTC	2460
Db	2401	CGAAGCGGGCGCACGATCTGAATATCGGCACATATTTGATCGTGCGCAAGCGGACAGGTC	2460
Qy	2461	CGTCTAACAGAGAAGTGGATCAGGTCGTAGATCGGGTTACGGAATTTAAAGACAGCTATG	2520
Db	2461	CGTCTAACAGAGAAGTGGATCAGGTCGTAGATCGGGTTACGGAATTTAAAGACAGCTATG	2520
Qy	2521	GACTGAAGATTGTGCAATGCTCTTGACCTGTTTGAAGCCAGACAGCGGGAAGCGCTCAAAG	2580
Db	2521	GACTGAAGATTGTGCAATGCTCTTGACCTGTTTGAAGCCAGACAGCGGGAAGCGCTCAAAG	2580
Qy	2581	ATGCAGGAGTAGACCGCTATAATCATAATTTGAATACGTACACAGAGAAACCAATTCAAACA	2640
Db	2581	ATGCAGGAGTAGACCGCTATAATCATAATTTGAATACGTACACAGAGAAACCAATTCAAACA	2640
Qy	2641	TCACAACCTCACATACATACGATGACAGAGTCAATACGGTTGAATTCGCAAAAGATCCG	2700
Db	2641	TCACAACCTCACATACATACGATGACAGAGTCAATACGGTTGAATTCGCAAAAGATCCG	2700
Qy	2701	GGCTGCTCTCCGTGTTTCAAGCGCCATATCCGGATGAAGAGACGAAACAGATGTCAATG	2760
Db	2701	GGCTGCTCTCCGTGTTTCAAGCGCCATATCCGGATGAAGAGACGAAACAGATGTCAATG	2760
Qy	2761	ACATCGCCAAAGCTTGAAAGGCTCTTGACCGGATCCATTCCTGTGAAATTTTTTGCATG	2820
Db	2761	ACATCGCCAAAGCTTGAAAGGCTCTTGACCGGATCCATTCCTGTGAAATTTTTTGCATG	2820
Qy	2821	CAATTGATGACACGCCGTTAGAGGCGCTCAACGNAATTAACCCGCTGTATTTGTTTAAAG	2880
Db	2821	CAATTGATGACACGCCGTTAGAGGCGCTCAACGNAATTAACCCGCTGTATTTGTTTAAAG	2880
Qy	2881	TGCTGGCGCTGTTCCGTTTTATCAATCCATCAAAAGAAATTCGCAATTCGCGAGGAAGAG	2940
Db	2881	TGCTGGCGCTGTTCCGTTTTATCAATCCATCAAAAGAAATTCGCAATTCGCGAGGAAGAG	2940
Qy	2941	AGGTCAATCTCCGCACATTCAGCCATTTAGGGCTTTACGCCGAAACTCCATTTTTGTGCG	3000
Db	2941	AGGTCAATCTCCGCACATTCAGCCATTTAGGGCTTTACGCCGAAACTCCATTTTTGTGCG	3000
Qy	3001	GAGACTACTTTAAACACTGCCGGCAGAGACGAGAGATCAATAATCTCTAGTGATT	3060
Db	3001	GAGACTACTTTAAACACTGCCGGCAGAGAGACGAGAGATCAATAATCTCTAGTGATT	3060
Qy	3061	TAGGCTTTGAAGTTGAATCAGTCAAGAAATGAAGGCTAGTTTTAAGTGCAGAAAGCTGAA	3120
Db	3061	TAGGCTTTGAAGTTGAATCAGTCAAGAAATGAAGGCTAGTTTTAAGTGCAGAAAGCTGAA	3120
Qy	3121	AGAATCAATAAAGCAATCGGTATGATGTGCAATTC	3156
Db	3121	AGAATCAATAAAGCAATCGGTATGATGTGCAATTC	3156

RESULT 2					
AR435692					
LOCUS	AR435692		3156 bp	DNA	linear
DEFINITION	Sequence 1 from patent US 6556721.				
ACCESSION	AR435692				
VERSION	AR435692.1		GI:40198766		

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 3156)
AUTHORS	Hohmann,H.-P., Mouncey,N.J., Schlieker,H.W. and Stebbins,J.W.
TITLE	Polynucleotide portions of the biofin operon from B. subtilis for use in enhanced fermentation
JOURNAL	Patent: US 6656721-A 1 02-DEC-2003; Roche Vitamins, Inc.; Parsippany, NJ
FEATURES	Location/Qualifiers
source	1..3156 /organism="unknown" /mol_type="genomic DNA"
ORIGIN	
Query Match	100.0%; Score 3156; DB 6; Length 3156;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3156; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GGATCCACGAGGTTTACGAGCCTTGAAAGATTGAATTCCTGGTTAAACGAGCGGTTAGACAGA 60
DB	1 GGATCCACGAGGTTTACGAGCCTTGAAAGATTGAATTCCTGGTTAAACGAGCGGTTAGACAGA 60
QY	61 ATGAAAGAACGCCGCCTGTAATCGCTAACCCTCGCGTCAATGGATGAGAGCGCCGGTTCAGAG 120
DB	61 ATGAAAGAACGCCGCCTGTAATCGCTAACCCTCGCGTCAATGGATGAGAGCGCCGGTTCAGAG 120
QY	121 AGGAATATTGATGGCGAATAATCAAACGGTCTGGTCTCTCAACAATATTATTAGGGCTCGCA 180
DB	121 AGGAATATTGATGGCGAATAATCAAACGGTCTGGTCTCTCAACAATATTATTAGGGCTCGCA 180
QY	181 AGCATAGACGTTTTGATCGATGCGCCCAACACAGCATTTGCAGCAATTTGGGACAGGAAGC 240
DB	181 AGCATAGACGTTTTGATCGATGCGCCCAACACAGCATTTGCAGCAATTTGGGACAGGAAGC 240
QY	241 AGCGGTTCAAGTTTAAACGACAGGCAATTCGGTCTGGCATGAAAAGCTAGAAAAAGAAGATT 300
DB	241 AGCGGTTCAAGTTTAAACGACAGGCAATTCGGTCTGGCATGAAAAGCTAGAAAAAGAAGATT 300
QY	301 GCCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTTCGAGCGGTTACTTGGCCAATGTC 360
DB	301 GCCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTTCGAGCGGTTACTTGGCCAATGTC 360
QY	361 GGTGTCTCTTTTCATCTCCAGAAAAGAAAGATGTCAATTTTAAAGTACACAGCTCAATCAT 420
DB	361 GGTGTCTCTTTTCATCTCCAGAAAAGAAAGATGTCAATTTTAAAGTACACAGCTCAATCAT 420
QY	421 GCAGATATGATCGACGGCTCCGACTTTTAAAGGCTGATACAGTTGTTTATCGGCATATT 480
DB	421 GCAGATATGATCGACGGCTCCGACTTTTAAAGGCTGATACAGTTGTTTATCGGCATATT 480
QY	481 GATATGATGATCTTTGAAAACAGCTGAATGAAACACAGCGTTATACAGCCCGTTTTATC 540
DB	481 GATATGATGATCTTTGAAAACAGCTGAATGAAACACAGCGTTATACAGCCCGTTTTATC 540
QY	541 GTAACACACGAGTATTTCAGCATGGATGGCACAAATCGCCCTCTTTGATCAGATCATCTCA 600
DB	541 GTAACACGAGATTTTCAGCATGGATGGCACAAATCGCCCTCTTTGATCAGATCATCTCA 600
QY	601 CTTGCGAAACGCTATCATGCTTCCTGGTGGTGAATGATCCCCACGAAACAGGAGTTTGG 660
DB	601 CTTGCGAAACGCTATCATGCTTCCTGGTGGTGAATGATCCCCACGAAACAGGAGTTTGG 660
QY	661 GGCATTTGGGACNAGGAACGAGTGAATACCTTTGGTGGTGTGTCGCCGACATTTGATCGGC 720
DB	661 GGCATTTGGGACNAGGAACGAGTGAATACCTTTGGTGGTGTGTCGCCGACATTTGATCGGC 720
QY	721 ACCTTTAAGCAAAAGCTGTTTGGCGGGAAGGAGTTTTTCGCGCAGGATCAGCGGCTCTTCATC 780
DB	721 ACCTTTAAGCAAAAGCTGTTTGGCGGGAAGGAGTTTTTCGCGCAGGATCAGCGGCTCTTCATC 780
QY	781 GACTTTTTTGCTGAACCATGCCAGAACATTTATCTTTTCAAAACCGCTATTTCGCCACCGCAGC 840

Db 781 |||||GAC|TTTTG|CTG|NA|CCAT|G|CC|AGA|ACAT|TTAT|CT|TTT|CA|AA|CG|CTA|TT|CC|G|CC|AG|CC|AG|C 840

Qy 841 TGTGGGCTGCTCAGAGGCTTTCAACATCA|TTGA|GCC|AG|CG|AG|G|A|A|A|A|CG|AC|AG|CTT 900

Db 841 TGTGGGCTGCTCAGAGGCTTTCAACATCA|TTGA|GCC|AG|CG|AG|G|A|A|A|A|CG|AC|AG|CTT 900

Qy 901 TTATTTTCTTATATCAGCATGATCAGAAC|CAGT|CT|GA|AG|AA|TAT|GGG|TTAT|GT|G|GT|G|AAA 960

Db 901 TTA|TTTT|CTT|AT|AT|CAG|CAT|GAT|C|GA|A|CC|AGT|CT|GA|AG|AA|TAT|GGG|TTAT|GT|G|GT|G|AAA 960

Qy 961 GGAGATCACACCGATTA|TT|CCT|GT|AGT|CAT|TTGG|CGAT|GCC|AT|AA|A|CG|T|CCTA|TTT 1020

Db 961 GGAGATCACACCGATTA|TT|CCT|GT|AGT|CAT|TTGG|CGAT|GCC|AT|AA|A|CG|T|CCTA|TTT 1020

Qy 1021 GCTG|AA|A|A|C|TG|C|AG|GG|CA|AG|G|AA|TTAT|G|CT|CT|G|CC|AT|TT|CG|CC|CG|CC|A|A|CG|TT|G|CG 1080

Db 1021 GCTG|AA|A|A|C|TG|C|AG|GG|CA|AG|G|AA|TTAT|G|CT|CT|G|CC|AT|TT|CG|CC|CG|CC|A|A|CG|TT|G|CG 1080

Qy 1081 CCGG|GT|GA|A|G|CC|G|AT|TC|GA|AG|CT|TGG|C|AG|C|AG|GT|CG|AG|AT|CAG|G|AA|T|GAG|TT|TATA 1140

Db 1081 CCGG|GT|GA|A|G|CC|G|AT|TC|GA|AG|CT|TGG|C|AG|C|AG|GT|CG|AG|AT|CAG|G|AA|T|GAG|TT|TATA 1140

Qy 1141 AAAT|AA|A|A|A|A|A|G|C|C|T|G|AA|A|G|G|T|G|T|TTTT|T|G|AT|G|G|TTT|G|AA|C|T|G|T|CT|TT|CTT 1200

Db 1141 AAAT|AA|A|A|A|A|A|G|C|C|T|G|AA|A|G|G|T|G|T|TTTT|T|G|AT|G|G|TTT|G|AA|C|T|G|T|CT|TT|CTT 1200

Qy 1201 ATCT|TG|AT|AC|AT|AT|AG|AA|AT|A|AC|GT|CA|TTTT|T|AT|TTTT|T|AT|TTTT|T|AG|T|G|CT|G|AA|A|G|G|T|G|CG 1260

Db 1201 ATCT|TG|AT|AC|AT|AT|AG|AA|AT|A|AC|GT|CA|TTTT|T|AT|TTTT|T|AT|TTTT|T|AG|T|G|CT|G|AA|A|G|G|T|G|CG 1260

Qy 1261 TTGA|GT|G|G|T|G|G|T|AT|G|T|G|T|TT|T|AA|G|T|AT|T|G|A|A|A|C|C|T|T|A|A|A|T|T|G|G|T|G|C|A|G 1320

Db 1261 TTGA|GT|G|G|T|G|G|T|AT|G|T|G|T|TT|T|AA|G|T|AT|T|G|A|A|A|C|C|T|T|A|A|A|T|T|G|G|T|G|C|A|G 1320

Qy 1321 AAAA|A|C|C|C|A|C|T|G|T|T|AA|G|T|T|A|A|G|T|G|A|C|T|A|A|A|C|A|A|A|T|A|C|T|A|A|A|T|G|A|G|G|G|T|T 1380

Db 1321 AAAA|A|C|C|C|A|C|T|G|T|T|AA|G|T|T|A|A|G|T|G|A|C|T|A|A|A|C|A|A|A|T|A|C|T|A|A|A|T|G|A|G|G|G|T|T 1380

Qy 1381 CTTT|TA|AT|T|AT|G|T|G|T|C|CT|A|AT|A|G|T|A|G|T|T|AT|T|C|A|G|AT|G|A|A|A|A|T|C|A|A|G|G|T|T|T|AG 1440

Db 1381 CTTT|TA|AT|T|AT|G|T|G|T|C|CT|A|AT|A|G|T|A|G|T|T|AT|T|C|A|G|AT|G|A|A|A|A|T|C|A|A|G|G|T|T|T|AG 1440

Qy 1441 TGG|A|C|A|G|A|A|A|A|G|T|G|G|A|A|G|T|G|A|C|C|A|T|G|T|T|T|G|A|A|A|G|C|A|G|G|T|T|T|A|A|G 1500

Db 1441 TGG|A|C|A|G|A|A|A|A|G|T|G|G|A|A|G|T|G|A|C|C|A|T|G|T|T|T|G|A|A|A|G|C|A|G|G|T|T|T|A|A|G 1500

Qy 1501 CTGA|TA|A|A|G|A|A|C|G|G|T|G|CT|C|C|A|A|AT|T|T|CT|T|AT|T|T|G|A|A|A|G|C|A|A|AT|CT|A|A|A|AT|T|AT|CT 1560

Db 1501 CTGA|TA|A|A|G|A|A|C|G|G|T|G|CT|C|C|A|A|AT|T|T|CT|T|AT|T|T|G|A|A|A|G|C|A|A|AT|CT|A|A|A|AT|T|AT|CT 1560

Qy 1561 GAA|A|G|G|A|AT|G|A|G|A|AT|G|T|G|A|AT|G|G|A|C|C|A|T|A|T|A|AT|G|A|T|G|A|G|A|A|A|G|A|AT|G|A|G 1620

Db 1561 GAA|A|G|G|A|AT|G|A|G|A|AT|G|T|G|A|AT|G|G|A|C|C|A|T|A|T|A|AT|G|A|T|G|A|G|A|A|A|G|A|AT|G|A|G 1620

Qy 1621 ATT|G|T|T|C|AT|G|A|A|T|T|A|A|G|G|A|C|G|A|T|T|G|A|T|A|A|T|G|G|G|AT|G|AT|G|T|T|A|A|G|C|T|A|TT 1680

Db 1621 ATT|G|T|T|C|AT|G|A|A|T|T|A|A|G|G|A|C|G|A|T|T|G|A|T|A|A|T|G|G|G|AT|G|AT|G|T|T|A|A|G|C|T|A|TT 1680

Qy 1681 GGT|G|T|T|AT|G|G|T|C|T|CT|T|G|G|T|C|G|T|C|A|G|A|C|T|G|AT|G|G|G|C|C|T|A|TT|C|G|G|AT|T|T|G|A|G|AT|G 1740

Db 1681 GGT|G|T|T|AT|G|G|T|C|T|CT|T|G|G|T|C|G|T|C|A|G|A|C|T|G|AT|G|G|G|C|C|T|A|TT|C|G|G|AT|T|T|G|A|G|AT|G 1740

Qy 1741 TGT|G|T|G|T|C|A|A|C|A|G|A|G|A|A|G|C|A|G|A|G|T|T|C|A|G|C|C|A|T|G|A|AT|G|A|A|T|G|A|G|G|T|G|G|A|G 1800

Db 1741 TGT|G|T|G|T|C|A|A|C|A|G|A|G|A|A|G|C|A|G|A|G|T|T|C|A|G|C|C|A|T|G|A|AT|G|A|A|T|G|A|G|G|T|G|G|A|G 1800

Qy 1801 GTG|G|A|G|T|G|A|T|T|T|G|A|T|G|A|G|A|G|A|T|T|C|T|A|G|A|T|T|G|C|AT|C|T|C|A|G|G|T|G|G|A|T|C|A 1860

Db 1801 GTG|G|A|G|T|G|A|T|T|T|G|A|T|G|A|G|A|G|A|T|T|C|T|A|G|A|T|T|G|C|AT|C|T|C|A|G|G|T|G|G|A|T|C|A 1860

Qy 1861 GATT|G|G|C|G|C|T|T|A|C|A|C|A|T|G|T|CA|AT|TTT|T|CT|CA|TTT|T|G|C|G|C|A|T|T|AT|G|A|T|T|G|C|G 1920

Db 1861 GATT|G|C|G|C|G|C|T|T|A|C|A|C|A|T|G|T|CA|AT|TTT|T|CT|AT|TTT|T|G|C|G|A|T|T|AT|G|A|T|T|C|A|G|G|T|G|G|A 1920

Qy 1921 TACT|TA|G|A|G|A|A|G|T|G|T|AT|C|A|A|C|T|G|T|T|A|A|AT|C|G|T|A|G|A|G|C|C|C|A|A|A|C|G|T|T|C|C|A|C|A|G|AT|G|C|G 1980

Db 1921 TACT|TA|G|A|G|A|A|G|T|G|T|AT|C|A|A|C|T|G|T|T|A|A|AT|C|G|T|A|G|A|G|C|C|C|A|A|A|C|G|T|T|C|C|A|C|A|G|AT|G|C|G 1980

Qy 1981 ATTT|G|T|G|C|C|CT|T|AT|C|G|T|A|G|A|G|A|G|C|T|G|T|T|T|G|A|AT|T|AT|C|A|G|C|G|C|A|A|AT|G|C|G|T|T|A|A|T|T|C|G|T 2040

Db 1981 ATTT|G|T|G|C|C|CT|T|AT|C|G|T|A|G|A|G|A|G|C|T|G|T|T|T|G|A|AT|T|AT|C|A|G|C|G|C|A|A|AT|G|C|G|T|T|A|A|T|T|C|G|T 2040

Qy 2041 GTG|C|A|A|G|G|A|C|C|G|A|C|A|A|C|AT|TT|T|C|A|C|AT|C|T|T|G|A|T|C|G|T|A|C|A|G|T|A|G|C|A|A|T|G|G|C|A|A|T|T|C|G|T 2100

Db 2041 GTG|C|A|A|G|G|A|C|C|G|A|C|A|A|C|AT|TT|T|C|A|C|AT|C|T|T|G|A|T|C|G|T|A|C|A|G|T|A|G|C|A|A|T|T|C|G|C|A|G|T|G|C|C 2100

Qy 2101 ATG|T|T|G|A|T|T|G|G|T|C|T|G|C|AT|C|AT|C|G|C|AT|C|G|T|T|A|T|A|C|G|A|C|G|A|G|C|G|T|C|G|G|T|C|T|T|A|A|C|T|G|A|A 2160

Db 2101 ATG|T|T|G|A|T|T|G|G|T|C|T|G|C|AT|C|AT|C|G|C|AT|C|G|T|T|A|T|A|C|G|A|C|G|A|G|C|G|T|C|G|G|T|C|T|T|A|A|C|T|G|A|A 2160

Qy 2161 GCAG|TT|A|A|G|C|A|A|T|C|A|G|AT|C|T|T|C|T|T|C|A|G|G|T|T|AT|G|A|C|C|AT|C|T|G|T|C|C|A|G|T|T|C|G|T|A|AT|G|T|CT 2220

Db 2161 GCAG|TT|A|A|G|C|A|A|T|C|A|G|AT|C|T|T|C|T|T|C|A|G|G|T|T|AT|G|A|C|C|AT|C|T|G|T|C|C|A|G|T|T|C|G|T|A|AT|G|T|CT 2220

Qy 2221 GGT|C|A|A|C|T|T|T|C|C|G|A|C|T|C|T|G|A|A|A|C|T|T|C|T|G|A|A|T|C|G|C|T|A|G|A|A|T|T|T|C|T|G|A|A|T|T|G|G|G|AT|T 2280

Db 2221 GGT|C|A|A|C|T|T|T|C|C|G|A|C|T|C|T|G|A|A|A|C|T|T|C|T|G|A|A|T|T|T|C|G|C|T|A|G|A|A|T|T|T|C|T|G|A|A|T|T|G|G|G|AT|T 2280

Qy 2281 CAGG|AG|T|G|G|A|C|A|G|A|A|C|G|A|C|A|C|G|G|AT|AT|A|T|A|G|T|G|G|AT|G|T|G|T|C|A|A|A|C|C|G|A|C|C|AT|T|T|G|A 2340

Db 2281 CAGG|AG|T|G|G|A|C|A|G|A|A|C|G|A|C|A|C|G|G|AT|AT|A|T|A|G|T|G|G|AT|G|T|G|T|C|A|A|A|C|C|G|A|C|C|AT|T|T|G|A 2340

Qy 2341 ATT|C|G|A|A|G|C|C|G|C|G|AT|T|G|AG|T|C|T|T|A|C|C|G|AT|G|G|T|G|A|A|T|A|A|G|G|A|A|A|C|G|T|C|T|T|C|A|A|G|G|C|G 2400

Db 2341 ATT|C|G|A|A|G|C|C|G|C|G|AT|T|G|AG|T|C|T|T|A|C|C|G|AT|G|G|T|G|A|A|T|A|A|G|G|A|A|A|C|G|T|C|T|T|C|A|A|G|G|C|G 2400

Qy 2401 CGA|A|G|C|G|G|C|G|C|A|C|G|A|T|C|T|G|A|A|T|A|T|C|G|G|C|A|C|A|T|T|G|T|AT|C|G|T|G|G|C|A|A|G|C|G|C|A|G|A|G|T|C 2460

Db 2401 CGA|A|G|C|G|G|C|G|C|A|C|G|A|T|C|T|G|A|A|T|A|T|C|G|G|C|A|C|A|T|T|G|T|AT|C|G|T|G|G|C|A|A|G|C|G|C|A|G|A|G|T|C 2460

Qy 2461 CGT|C|T|A|A|C|A|G|A|A|G|T|G|G|AT|C|A|G|G|T|C|G|T|A|G|T|G|G|G|T|T|C|A|G|G|A|A|T|T|A|A|G|A|G|A|C|G|T|AT|G 2520

Db 2461 CGT|C|T|A|A|C|A|G|A|A|G|T|G|G|AT|C|A|G|G|T|C|G|T|A|G|T|G|G|G|T|T|C|A|G|G|A|A|T|T|A|A|G|A|G|A|C|G|T|AT|G 2520

Qy 2521 GACT|G|A|A|G|A|T|T|G|T|G|C|A|T|G|T|T|G|G|A|C|T|G|T|T|G|A|A|G|C|A|G|A|G|C|G|G|C|T|C|A|A|A|G 2580

Db 2521 GACT|G|A|A|G|A|T|T|G|T|G|C|A|T|G|T|T|G|G|A|C|T|G|T|T|G|A|A|G|C|A|G|A|G|C|G|G|C|T|C|A|A|A|G 2580

Qy 2581 ATG|C|A|G|A|G|T|A|G|A|C|C|G|C|T|A|T|A|A|T|T|G|A|A|T|A|C|G|T|C|A|G|A|G|A|A|A|C|C|A|T|T|C|A|A|A|C|A 2640

Db 2581 ATG|C|A|G|A|G|T|A|G|A|C|C|G|C|T|A|T|A|A|T|T|G|A|A|T|A|C|G|T|C|A|G|A|G|A|A|A|C|C|A|T|T|C|A|A|A|C|A 2640

Qy 2641 TCA|C|A|A|C|C|T|C|A|C|A|T|A|C|A|T|A|C|G|A|T|G|A|C|A|G|T|C|A|A|T|A|C|G|G|T|T|G|A|A|T|C|G|A|A|A|G|A|A|T|C|G|G 2700

Db 2641 TCA|C|A|A|C|C|T|C|A|C|A|T|A|C|A|T|A|C|G|A|T|G|A|C|A|G|T|C|A|A|T|A|C|G|G|T|T|G|A|A|T|C|G|A|A|A|G|A|A|T|C|G|G 2700

Qy 2701 GGT|G|T|G|T|C|C|G|T|T|C|A|G|G|C|C|A|T|T|AT|C|G|G|AT|G|A|A|G|A|G|A|C|G|A|A|A|C|A|G|A|G|AT|G|C|A|T|T|G 2760

Db 2701 GGT|G|T|G|T|C|C|G|T|T|C|A|G|G|C|C|A|T|T|AT|C|G|G|AT|G|A|A|G|A|G|A|C|G|A|A|A|C|A|G|A|G|AT|G|C|A|T|T|G 2760

Qy 2761 ACAT|C|G|C|A|A|A|G|G|T|T|G|A|A|G|G|C|T|T|T|G|A|C|G|G|A|T|T|C|C|AT|T|C|T|G|T|G|A|A|T|T|T|T|T|G|C|A|T|G 2820

Db 2761 ACAT|C|G|C|A|A|A|G|G|T|T|G|A|A|G|G|C|T|T|T|G|A|C|G|G|A|T|T|C|C|AT|T|C|T|G|T|G|A|A|T|T|T|T|T|G|C|A|T|G 2820

Qy 2821 CAAT|T|G|A|T|G|C|A|C|G|C|G|T|T|A|G|A|G|G|C|G|T|C|A|C|G|A|A|T|T|A|A|A|C|C|C|G|C|T|G|T|A|T|T|G|T|T|T|A|A|A|G 2880

Db 2821 CAAT|T|G|A|T|G|C|A|C|G|C|G|T|T|A|G|A|G|G|C|G|T|C|A|C|G|A|A|T|T|A|A|A|C|C|C|G|C|T|G|T|A|T|T|G|T|T|T|A|A|A|G 2880

Qy 2881 TGCT|G|G|G|C|G|T|T|G|T|T|C|G|T|T|T|AT|C|A|AT|C|C|AT|C|A|A|A|G|A|A|A|T|T|C|G|C|A|T|T|T|C|C|G|G|A|G|A|A|G|A|G 2940

Db 2881 TGCT|G|G|G|C|G|T|T|G|T|T|C|G|T|T|T|AT|C|A|AT|C|C|AT|C|A|A|A|G|A|A|A|T|T|C|G|C|A|T|T|T|C|C|G|G|A|G|A|A|G|A|G 2940

Qy 2941 AGGT|C|A|A|T|C|T|C|C|G|C|A|C|A|T|T|G|C|A|G|C|A|T|T|A|G|G|G|C|T|T|T|A|C|G|C|G|C|A|A|A|C|T|C|A|A|T|T|T|T|G|T|G|C 3000

Db 2941 AGGT|C|A|A|T|C|T|C|C|G|C|A|C|A|T|T|G|C|A|G|C|A|T|T|A|G|G|G|C|T|T|T|A|C|G|C|G|C|A|A|A|C|T|C|C|A|A|T|T|T|T|G|T|G|C 3000

1561 GAAAGGGAATGAGAAATAGTGAATGACCAATTAATATGACTAGAGAAAGAAATGAAG 1620 Db
1621 ATTGTTCAATGAAATTAAGGAACGAATATTGATTAATATGCGGATGATGTTAAGGCTATT 1680 Qy
1621 ATTGTTCAATGAAATTAAGGAACGAATATTGATTAATATGCGGATGATGTTAAGGCTATT 1680 Db
1681 GGTGTTATGCTCTCTTGGTCTGTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATG 1740 Qy
1681 GGTGTTATGCTCTCTTGGTCTGTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATG 1740 Db
1741 TGTGTCATGTCACAGAGGAAGAGAGTTTCAGCCATGAATGGACAACCGGTGAGTGAAG 1800 Qy
1741 TGTGTCATGTCACAGAGGAAGAGAGTTTCAGCCATGAATGGACAACCGGTGAGTGAAG 1800 Db
1801 GTGAAAGTGAATTTTGATAGCGAAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCA 1860 Qy
1801 GTGAAAGTGAATTTTGATAGCGAAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCA 1860 Db
1861 GATTGGCCGCTTACACATGGTCAATTTTCTATTTTCCGATTTATGATTTCAAGGTGGA 1920 Qy
1861 GATTGGCCGCTTACACATGGTCAATTTTCTATTTTCCGATTTATGATTTCAAGGTGGA 1920 Db
1921 TACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAGGCCCAAAAGTTTCCACGATGCG 1980 Qy
1921 TACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAGGCCCAAAAGTTTCCACGATGCG 1980 Db
1981 ATTTGTGCCCCATTGTAAGAGAGTGTTCGAATATGCAAGCAAAATGCGTAAATTTTGGT 2040 Qy
1981 ATTTGTGCCCCATTGTAAGAGAGTGTTCGAATATGCAAGCAAAATGCGTAAATTTTGGT 2040 Db
2041 GTCAAGGACCGACAACATTTCTACCATCTTGTACTGTACAGGTAGCAATGGCAGGTGCC 2100 Qy
2041 GTCAAGGACCGACAACATTTCTACCATCTTGTACTGTACAGGTAGCAATGGCAGGTGCC 2100 Db
2101 ATGTTGATTTGCTGTCATCATCGCATCTGTTATATACGACGAGCGCTTCGGTCTTAACGAA 2160 Qy
2101 ATGTTGATTTGCTGTCATCATCGCATCTGTTATATACGACGAGCGCTTCGGTCTTAACGAA 2160 Db
2161 GCAGTTAAGCAATCGATCTTCTTCAGGTTATGACCATCTGTGCCAGTTTCGTAATGTCT 2220 Qy
2161 GCAGTTAAGCAATCGATCTTCTTCAGGTTATGACCATCTGTGCCAGTTTCGTAATGTCT 2220 Db
2221 GGTCAACTTTCCGACTCTGAGAACTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280 Qy
2221 GGTCAACTTTCCGACTCTGAGAACTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280 Db
2281 CAGGAGTGGACAGAACCGACGGATATATAGTGGATGTGTCAAAACGATACCATTTTGA 2340 Qy
2281 CAGGAGTGGACAGAACCGACGGATATATAGTGGATGTGTCAAAACGATACCATTTTGA 2340 Db
2341 ATTCGAAGGCCGATTTGAGTCTTACCGGATGTTGAATAGGAACCGTCTGAGGCG 2400 Qy
2341 ATTCGAAGGCCGATTTGAGTCTTACCGGATGTTGAATAGGAACCGTCTGAGGCG 2400 Db
2401 CGAAGCGGCGCACCATCTGAATATCGGCACATATTGTATCGTGGCAAGCGGACAGAGTC 2460 Qy
2401 CGAAGCGGCGCACCATCTGAATATCGGCACATATTGTATCGTGGCAAGCGGACAGAGTC 2460 Db
2461 CGTCTAAACAGAGAAGTGGATCAGTCTGTAGATGCGGTTTCAGGAAATTAAGAGACGTATG 2520 Qy
2461 CGTCTAAACAGAGAAGTGGATCAGTCTGTAGATGCGGTTTCAGGAAATTAAGAGACGTATG 2520 Db
2521 GACTGAAGATTTGTGCAATCTTGGACTGTTGAAGCCAGACAGCGGAGCGGCTCAAG 2580 Qy
2521 GACTGAAGATTTGTGCAATCTTGGACTGTTGAAGCCAGACAGCGGAGCGGCTCAAG 2580 Db
2581 ATGCAGGAGTAGACCGCTTAAATCAATTTTGAATACGTACAGAGAAACCATTTCAACA 2640 Qy
2581 ATGCAGGAGTAGACCGCTTAAATCAATTTTGAATACGTACAGAGAAACCATTTCAACA 2640 Db
2641 TCACAACCTCACAATACATACGATGACAGAGTCAATACCGGTTGAAATCGCAAAAAGATCGG 2700 Qy
2641 TCACAACCTCACAATACATACGATGACAGAGTCAATACCGGTTGAAATCGCAAAAAGATCGG 2700 Db

2701 GGCTGCTCCGTGTTCAAGCGCCATTATCGSGATGAGGAGACGAAACAGGATGTCATTG 2760 Qy
2701 GGCTGCTCCGTGTTCAAGCGCCATTATCGSGATGAGGAGACGAAACAGGATGTCATTG 2760 Db
2761 ACATCGCAAAAGGTTGAAGGCTCTTGACGCGGATTCATTTCTGTGAATTTTTTGCATG 2820 Qy
2761 ACATCGCAAAAGGTTGAAGGCTCTTGACGCGGATTCATTTCTGTGAATTTTTTGCATG 2820 Db
2821 CAATTGATGCGACGCCGTTAGAGGCGTCAACGAATTTAAACCCGCTGATTTTAAAG 2880 Qy
2821 CAATTGATGCGACGCCGTTAGAGGCGTCAACGAATTTAAACCCGCTGATTTTAAAG 2880 Db
2881 TGCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGATTTTCGGAGGAAGAG 2940 Qy
2881 TGCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGATTTTCGGAGGAAGAG 2940 Db
2941 AGGTCAATCTCCGCACATTCAGCAGCATTTAGGCGTTTACGCGCAAACTCCATTTTGTGCG 3000 Qy
2941 AGGTCAATCTCCGCACATTCAGCAGCATTTAGGCGTTTACGCGCAAACTCCATTTTGTGCG 3000 Db
3001 GAGACTACTTAACAACTGCGCGGCAAGAGGAGCGAGGATCAATAAATGCTGAGTGATT 3060 Qy
3001 GAGACTACTTAACAACTGCGCGGCAAGAGGAGCGAGGATCAATAAATGCTGAGTGATT 3060 Db
3061 TAGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTCGGAAGAGTGAA 3120 Qy
3061 TAGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTCGGAAGAGTGAA 3120 Db
3121 AGAATCAATAAAGCAATCGGTATGATGTCGAATTC 3156 Qy
3121 AGAATCAATAAAGCAATCGGTATGATGTCGAATTC 3156 Db

RESULT 4
AR173058
LOCUS AR173058 8478 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6303377.
ACCESSION AR173058
VERSION AR173058.1 GI:17912549
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8478)
AUTHORS Bower,S.Grant., Perkins,J.B., Yocum,R.Rogers. and Pero,J.G.
TITLE Biotin biosynthesis in Bacillus subtilis
JOURNAL Patent: US 6303377-A 1 16-OCT-2001;
FEATURES Location/Qualifiers
source 1..8478
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 51.1%; Score 1612.6; DB 6; Length 8478;
Best Local Similarity 73.0%; Pred. No. 0;
Matches 2303; Conservative 0; Mismatches 624; Indels 226; Gaps 9;
Qy 2 GATCCACGAGTTACGAGCTTGAAGATTGATTCCTCGTTAAACGAGCGGTTAGACAGAA 61
Db 2524 GATCCACGAGTTACGAGCTTGAAGATTGATTCCTCGTTAAACGAGCGGTTAGACAGAA 2583
Qy 62 TGAAGAAGCGCGGTACATCGTAACTCGGTTCAATGGATGGAGCGCGGTTCCAGAGA 121
Db 2584 TGAAGAAGCGCGGTACATCGTAACTCGGTTCAATGGATGGAGCGCGGTTCCAGAGA 2643
Qy 122 GGAATATTGATGGCGAAAAATCAAAACGGTCTCGTCTCTCAAAACAATTTATTTAGGGCTCGCAA 181
Db 2644 GGAATATTGATGGCGAAAAATCAAAACGGTCTCGTCTCTCAAAACAATTTATTTAGGGCTCGCAA 2703
Qy 182 GCGATAGACGTTGATCGATGCGAGCCCAACAGCAATTTGCGAGCAATTTGGGACAGGAAGCA 241
Db 2704 GCGATAGACGTTGATCGATGCGAGCCCAACAGCAATTTGCGAGCAATTTGGGACAGGAAGCA 2763

QY 242 GCGTTTCAAGTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATTG 301
DB 2764 GCGTTTCAAGTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATTG 2823
QY 302 CCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTCG 361
DB 2824 CCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTCG 2883
QY 362 GTGTCCTTTCATCTCCAGAAAAGGAAGATGTCATTTTAAAGTGACAGCTCAATCATG 421
DB 2884 GTGTCCTTTCATCTCCAGAAAAGGAAGATGTCATTTTAAAGTGACAGCTCAATCATG 2943
QY 422 CAAGTATGATCGAGCGCTCGCGACTTTCTAAGGCTGATACAGTGTGTTTATCGGCATATTG 481
DB 2944 CAAGTATGATCGAGCGCTCGCGACTTTCTAAGGCTGATACAGTGTGTTTATCGGCATATTG 3003
QY 482 ATATGAATGATCTTGAANAACAGCTGAATGAACACAGCGTTATCAGCGCGCTTTATCG 541
DB 3004 ATATGAATGATCTTGAANAACAGCTGAATGAACACAGCGTTATCAGCGCGCTTTATCG 3063
QY 542 TAAACAGACGAGTATTCAGCATGATGCGCAATTCGCCCTCTTGATCAGATCATCTCAC 601
DB 3064 TAAACAGACGAGTATTCAGCATGATGCGCAATTCGCCCTCTTGATCAGATCATCTCAC 3123
QY 602 TTGCGAAACGCTATCATGCCCTTCGTGTCGTGATGATGCCAGCAACAGGAGTTTGG 661
DB 3124 TTGCGAAACGCTATCATGCCCTTCGTGTCGTGATGATGCCAGCAACAGGAGTTTGG 3183
QY 662 GCGATTCGGGACAGGAAACAGATGAATACATTTGGTGTGTTGTCGCCGACATTTGTTATCGGCA 721
DB 3184 GCGATTCGGGACAGGAAACAGATGAATACATTTGGTGTGTTGTCGCCGACATTTGTTATCGGCA 3243
QY 722 CCTTAACCAAGCTGTGTCGCGGAGAGGTTTGGCGGAGATCAGCGCTCTTCATCG 781
DB 3244 CCTTAACCAAGCTGTGTCGCGGAGAGGTTTGGCGGAGATCAGCGCTCTTCATCG 3303
QY 782 ACTTTTGTGTAACCATGCCAGAACATTTATCTTTCAAAACCGCTATTTCCGCCAGCCAGCT 841
DB 3304 ACTTTTGTGTAACCATGCCAGAACATTTATCTTTCAAAACCGCTATTTCCGCCAGCCAGCT 3363
QY 842 GTGCGGTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGAGGAGGAAAACGACAGCTTT 901
DB 3364 GTGCGGTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGAGGAGGAAAACGACAGCTTT 3423
QY 902 TATTTTCTTATATCAGCATGATCAGAACCACTCTGGAAGATATGGGTTATGTGTGAAG 961
DB 3424 TATTTTCTTATATCAGCATGATCAGAACCACTCTGGAAGATATGGGTTATGTGTGAAG 3483
QY 962 GAGATCACACCGGATTTATCTGTAGTCAATTTGGCGATGCCATPAAAACCGTCTTATTTG 1021
DB 3484 GAGATCACACCGGATTTATCTGTAGTCAATTTGGCGATGCCATPAAAACCGTCTTATTTG 3543
QY 1022 CTGAAAACCTGAGGCAAGGAAATTTATGCTCTGCGCATTTGGCGCCCAACCGTTGCGC 1081
DB 3544 CTGAAAACCTGAGGCAAGGAAATTTATGCTCTGCGCATTTGGCGCCCAACCGTTGCGC 3603
QY 1082 CGGCTGAACCGCGATTCGAAGCTTTGGCGAGCAGGTCGAGATCAGGGAATGAGTTTATAA 1141
DB 3604 CGGCTGAACCGCGATTCGA-----ATTA 3627
QY 1142 AATAAAAAAGACCTGAAAAGGTGCTTTTTTTTGTGTTTGAACCTGTCTTTCTTCTTA 1201
DB 3628 CAATCAGCTCTGACCACAGATATGGGTGATATTGATCATTTGCTGCAACAACATTTCAATCAA 3687
QY 1202 TCTTGATACATATAGAAATAACGTCATTTTTTATTTTATTTTATTTTGTGCTGAAGGTCGCT 1261
DB 3688 TCGAAAAGGAGCTGCACATCATTTTGAAGGGTTTTTTTGTGACCGGAACTGA----- 3738
QY 1262 TGAAGTGTGTGATGTGTTTTTAAAGTATTCAAAACCCCTTAAANTTGGTTGCAACAGA 1321
DB 3739 -----TACAGAAGTAGGGAACCGGTTATATCCAGCGGTCTTG- 3776

QY 1322 AAAACCCCATCTGTTAAAGTTATAGTACTAAACAAATAAATAAGATGAGGTTTC 1381
DB 3777 -----CTGCCCTTATTGAAAAGACAATAATAGACATCTCGGGGTGTA 3816
QY 1382 TTTTAAATATTATGTCCTTAATAGTAGCATTTTATTAGATGAAAATCAAGGGTTTTAGT 1441
DB 3817 TAAACCATTTTTTAAAGCGGGATATCGCGCCATCATCCAGATAGTGATACAAGTTTTGCTGAA 3876
QY 1442 GGAACAAGCAAAAAGTGGAAAAGTGAGACCATGTGCTTAGAAGACGAGTTTATTAATAGC 1501
DB 3877 AGATATGTCGCAG----- 3890
QY 1502 TGAATAAGAACGGTGTCTCCAAATATTCTTATTTAGAAAAGCAAAATCTAAAATATTCTG 1561
DB 3891 -----CCAGTCTTCTCTCTGAAGACATTTACGCTTTTGCCTTTCAA 3930
QY 1562 AAAAGGAATGAGAAATAGTGAATGGACCAATATAATGACTAGAGAAGAAAGATGAAGA 1621
DB 3931 GGGCCGCTTTCACCATACGTTGACGGGAAACTTTGAGGGAAAGACTGTCCACATGGGAAGA 3990
QY 1622 TTGTTTCATGAAAATTAAGGAACGAAATATTGGATAAATAATGGGGATGATGTTTAAAGCTATTG 1681
DB 3991 GGTTTTAAGCCATTTGGGGCGGATTAGAGAAA-----CATGAATGCTTCATCGTAG 4043
QY 1682 GTGTTTATGGCTCTCTTTGGTCTGACATGATGGGCCCTTATTCGGATATTAGATGATGAT 1741
DB 4044 AAGTGCAGCGCGTATTCTGTGCCATTTGGGAGAGGACTATTTGGTCACTCATGTCTATAA 4103
QY 1742 GTGTCATGTCAAACAGAGGAGCAGAGTTTCAGCCATGAAATGGACAAACCGGTGAGTGAAGG 1801
DB 4104 -----AAGCTTGCAGCTTCCCATGATTTTGTGGCGCTCTCCGCTTTGGAACC 4153
QY 1802 TGAAGTGAATTTTGTATAGCGAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAG 1861
DB 4154 ATTAATCATACCTTTTAACTGTAATATGCAGAAAGCATGGGCTTCCAAATCGCCGGA 4213
QY 1862 ATTGGCGGCTTACATGTCATTTTTCTCTATTTTGGCCGATTTATGATTCAGGTGAGT 1921
DB 4214 ATTAATCATCAATGGAATCAGTGACTCTCTCTGATGAAGATGAAAACCAATCTCTGAGATG 4273
QY 1922 ACTTAGAAGAGTATCAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATCGGA 1981
DB 4274 ATTGAGCGCTTTATGCGGTGTGCCGATTTTAGGGGTTACGCCAAAGCTTGCACAGTG--- 4330
QY 1982 TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGACAGCAAAATGGCGTAATATTCTGTG 2041
DB 4331 -----ACGAAGAAACGGTCTCATATGTTGAAGACCATA----- 4367
QY 2042 TGCAAGGACCGACAAATTTCTACCATCTTGTACTGTACAGGTAGCAATGGCAGGTGCCA 2101
DB 4368 -----TCAATCTATCATTTACTGATGAATCAAGTGGGGTATGA 4405
QY 2102 TGTGATTTGTCATCATCGCATCTGTTTATACGAGAGCGCTTCGGTCTTAACTGAAG 2161
DB 4406 GAATGAATCAATGATGGAATCTCGAGACCGGGTGTGGCTGGAGCAAGTGAAGTGAAG 4465
QY 2162 CAGTTAAGCAATCAGATCTTCTTTAGGTTATACCACTCTGTGCCAGTTCGTATGTCG 2221
DB 4466 AAGAGGCGCTTCAATATTACATTTGCTGATGAAGATATTTTGTCTATTATGACCGGG 4525
QY 2222 GTCAACTTTCCGACTCTGAGAAACCTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTC 2281
DB 4526 CTTTTTACATCAGAAAACACTTTTACGAAAAGGATAAGAGCTCAATATGATTAATG 4585
QY 2282 AGAGTTGGACAGAACGACACGGATATATAGTGGATGTCTCAAAACGCAATACCAATTTGAA 2341
DB 4586 CGAAATCCCGGGCTCTGCGCCCGAAAACCTGCGGCTATTGTTTCA-----GTCTCGCAT 4637
QY 2342 TTCGAAAAGCGCCGATTTAGTCTTTACCGGATGTTGAATAAGGAAACGCTGCTTGAAGCGC 2401
DB 4638 TTCGAAAAGCGCCGATTTAGTCTTTACCGGATGTTGAATAAGGAAACGCTGCTTGAAGCGC 4697
QY 2402 GAAGCGGGCGCACGATCTGAATATCGGCAATATTGTATCGTGGCAAGCGGACAGAGTCC 2461

Db	4698	GAAGCGGCGCACGATCTGAATATATCGGCACATATTGTATGTGGCAAGCGCGAGGTCC	4757	
Qy	2462	GTCTAAACAGAGAGTGGATCAGGTCTAGATGCGGTTCAGAGAAATTAAAGAGAGTATGG	2521	
Db	4758	GTCTAAACAGAGAGTGGATCAGGTCTAGATGCGGTTCAGAGAAATTAAAGAGAGTATGG	4817	
Qy	2522	ACTGAAGATTTGTGCATGCTCTTGGACTCTTGAAGCCAGAGCAGGCGAAGCGGCTCAAGA	2581	
Db	4818	ACTGAAGATTTGTGCATGCTCTTGGACTCTTGAAGCCAGAGCAGGCGAAGCGGCTCAAGA	4877	
Qy	2582	TGCAGGAGTAGACCGCTATAATCAATAATTGAATACGTCACAGAGAAACCAATTCAAAACAT	2641	
Db	4878	TGCAGGAGTAGACCGCTATAATCAATAATTGAATACGTCACAGAGAAACCAATTCAAAACAT	4937	
Qy	2642	CACAACTTCACATACATACGATGACAGATCAATACGGTTGAAATCGCAAAAGATCGGG	2701	
Db	4938	CACAACTTCACATACATACGATGACAGATCAATACGGTTGAAATCGCAAAAGATCGGG	4997	
Qy	2702	GCTGTCTCGGTTCAGCGCGCATTTATCGGATGAAGAGAGACGAAACAGGATGTCATTGA	2761	
Db	4998	GCTGTCTCGGTTCAGCGCGCATTTATCGGATGAAGAGAGACGAAACAGGATGTCATTGA	5057	
Qy	2762	CATGCCAAAAAGCTTGAAGGCTTTGACCGCGATTTCCATTCTCTGTGAATTTTTTGCATGC	2821	
Db	5058	CATGCCAAAAAGCTTGAAGGCTTTGACCGCGATTTCCATTCTCTGTGAATTTTTTGCATGC	5117	
Qy	2822	AATTGATGGCAGCGGTTAGAGGCGTCAACGAAATTAACCGCGCTGATTTGTTAAAGT	2881	
Db	5118	AATTGATGGCAGCGGTTAGAGGCGTCAACGAAATTAACCGCGCTGATTTGTTAAAGT	5177	
Qy	2882	GCTGGCGTGTTCGGTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAAAGAGA	2941	
Db	5178	GCTGGCGTGTTCGGTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAAAGAGA	5237	
Qy	2942	GGTCAATCTCCGCACATTCAGACCAATTAGGCGTTTACGCCCAACTCCATTTTGTGCGG	3001	
Db	5238	GGTCAATCTCCGCACATTCAGACCAATTAGGCGTTTACGCCCAACTCCATTTTGTGCGG	5297	
Qy	3002	AGACTACTTAACACTGCGGCGCAGAGAGACGAGGATCATAAATGCTGAGTGATTT	3061	
Db	5298	AGACTACTTAACACTGCGGCGCAGAGAGACGAGGATCATAAATGCTGAGTGATTT	5357	
Qy	3062	AGGCTTTCAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTGCAGAAAGCTGAAA	3121	
Db	5358	AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTGCAGAAAGCTGAAA	5417	
Qy	3122	GAATCAATAAAGCAATCGGTATGATGCGAAT	3154	
Db	5418	GAATCAATAAAGCAATCGGTATGATGCGAAT	5450	

RESULT 5	AR630545	8478 bp	DNA	linear	PAT 14-FEB-2005
LOCUS	Sequence 1 from patent US 6841366.				
DEFINITION	AR630545				
ACCESSION	AR630545				
VERSION	AR630545.1	GI:59765691			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 8478)				
AUTHORS	Bower, S.G., Perkins, J.B., Yocum, R.R. and Pero, J.G.				
TITLE	Biotin biosynthesis in bacillus subtilis				
JOURNAL	Patent: US 6841366-A 1 11-JAN-2005;				
	DSM IP Assets B.V.; Heerlen;				
	NLX;				
FEATURES	Location/Qualifiers				
source	1..8478				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					

Query Match	51.1%;	Score 1612.6;	DB 6;	Length 8478;	
Best Local Similarity	73.0%;	Pred. No. 0;			
Matches 2303;	Conservative	0;	Mismatches 624;	Indels 226;	Gaps 9;
Qy	2	GATCACGAGGTTACGAGCCTTGAAGATTGATCTCTGGTTAAACGAGCGGTTAGACAGAA	61		
Db	2524	GATCCACGAGGTTACGAGCCTTGAAGATTGATCTCTGGTTAAACGAGCGGTTAGACAGAA	2583		
Qy	62	TGAAGAAGCGCGGTACATCGTAACTGCGGTCAATGGATGGAGCGCGGTTCCAGAGA	121		
Db	2584	TGAAGAAGCGCGGTACATCGTAACTGCGGTCAATGGATGGAGCGCGGTTCCAGAGA	2643		
Qy	122	GGAATATTGATGGCGAAAAATCAAAACGGTCTGGTCTCTCAAAACAAATTATTATGGGCTCGCAA	181		
Db	2644	GGAATATTGATGGCGAAAAATCAAAACGGTCTGGTCTCTCAAAACAAATTATTATGGGCTCGCAA	2703		
Qy	182	GCATAGACGTTTGCATCGATGCGAGCCAAAACAGCATTTGCAGCAATTTGGGACAGGAAGCA	241		
Db	2704	GCATAGACGTTTGCATCGATGCGAGCCAAAACAGCATTTGCAGCAATTTGGGACAGGAAGCA	2763		
Qy	242	GCGGTTTCAAGTTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAGAGATTG	301		
Db	2764	GCGGTTTCAAGTTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAGAGATTG	2823		
Qy	302	CCAGCTTTAAACTGACAGAAAGCGCCCTGCTGTTTTCGAGCGGTTACTTTGGCCCAATGTCG	361		
Db	2824	CCAGCTTTAAACTGACAGAAAGCGCCCTGCTGTTTTCGAGCGGTTACTTTGGCCCAATGTCG	2883		
Qy	362	GTGTCCTTTTCACTCTTGCAGAAAAGGAAGATGTCATTTTAAAGTGACAGCTCAATCATG	421		
Db	2884	GTGTCCTTTTCACTCTTGCAGAAAAGGAAGATGTCATTTTAAAGTGACAGCTCAATCATG	2943		
Qy	422	CAAGTATGATCGAGCGGTCGCGACTTTCTAAGGCTGATACAGTTGTTTATCGGCATATTG	481		
Db	2944	CAAGTATGATCGAGCGGTCGCGACTTTCTAAGGCTGATACAGTTGTTTATCGGCATATTG	3003		
Qy	482	ATATGAATGATCTTGAATAAACAAGCTGAATGAAACAACAGCGGTTATCAGCGCGGTTTATCG	541		
Db	3004	ATATGAATGATCTTGAATAAACAAGCTGAATGAAACAACAGCGGTTATCAGCGCGGTTTATCG	3063		
Qy	542	TAAACAGCGGAGTATTACAGCATGGATGGCAAAATCGCCCTCTTTCATCAGATCATCTCAC	601		
Db	3064	TAAACAGCGGAGTATTACAGCATGGATGGCAAAATCGCCCTCTTTCATCAGATCATCTCAC	3123		
Qy	602	TTGCGAAACGCTATCATGCTTCGTCGTTGATGATGCCACGCAACAGAGTTTGG	661		
Db	3124	TTGCGAAACGCTATCATGCTTCGTCGTTGATGATGCCACGCAACAGAGTTTGG	3183		
Qy	662	GCATTCGCGGACCAAGGAACGAGTGAATACCTTTGGTGTGTTTGTCCGACATTTGTTATCGGCA	721		
Db	3184	GCATTCGCGGACCAAGGAACGAGTGAATACCTTTGGTGTGTTTGTCCGACATTTGTTATCGGCA	3243		
Qy	722	CCTTAAACAAAGCTTTGGCGCGGAAGAGGTTTTCGCGCAGGATCAGCGGCTTTCATCG	781		
Db	3244	CCTTAAACAAAGCTTTGGCGCGGAAGAGGTTTTCGCGCAGGATCAGCGGCTTTCATCG	3303		
Qy	782	ACTTTTTCGTCGAAACATGCCAGACATTTTATCTTTTCAAAACCGCTATTTCGCCAGCCAGCT	841		
Db	3304	ACTTTTTCGTCGAAACATGCCAGACATTTTATCTTTTCAAAACCGCTATTTCGCCAGCCAGCT	3363		
Qy	842	GTGCGGCTGCTCACGAGGCTTTTCAACATCATTTGAAGCCAGCAGGGAAGAAACGACAGCTTT	901		
Db	3364	GTGCGGCTGCTCACGAGGCTTTTCAACATCATTTGAAGCCAGCAGGGAAGAAACGACAGCTTT	3423		
Qy	902	TATTTTCTTATATCAGCATGATCAGAACAGCTCTGGAAGATATGGGTTATGTGGTGAAG	961		
Db	3424	TATTTTCTTATATCAGCATGATCAGAACAGCTCTGGAAGATATGGGTTATGTGGTGAAG	3483		
Qy	962	GAGATCACACACCGATTATTTCTGTAGTCATTTGGGATGCCCATATAAACGGTCTCTATTG	1021		
Db	3484	GAGATCACACACCGATTATTTCTGTAGTCATTTGGGATGCCCATATAAACGGTCTCTATTG	3543		

Qy	1022	CTGAAAACATGCGAGGGCAAGGGAATTTATGCTCTCTGCCATCTCGCGCCGCAACCGTTGCGC	1081
Db	3544	CTGAAAACATGCGAGGGCAAGGGAATTTATGCTCTCTGCCATCTCGCGCCGCAACCGTTGCGC	3603
Qy	1082	CGGGTGAAGCGCGGATTCGAAGCTTGGGCAGCAGGTGCGAGATCAGGGAAATGAGTTTATAA	1141
Db	3604	CGGGTGAAGCGCGGATTCGA	3627
Qy	1142	AATAAAAAAGCACCTGAAAAGGTGCTTTTTTTTGATGGTTTGAACCTGTCTTCTTCTTA	1201
Db	3628	CAATCAGCTCTGACCAACAGTATGGGTGATATGTATCATTTTGTCTGCAAAACATTTCAATCAA	3687
Qy	1202	TCTTGATACATATAGAATAACGTCATTTTATTTTATTTTATTTTATTTTGTCTGAAAGTCCGT	1261
Db	3688	TCGGAAGAGAGCTGCACATCATTTTGAAGGGTTTTTTTTGTGACGGAACTGA	3738
Qy	1262	TGAAGTCTTGATATGTTGTTTTTAAAGTATTGAAAACCTTAAAAATTTGTTTGCACAGA	1321
Db	3739	-----TACAGAAGTAGGNAACCGTTATATCCAGCGGCTTTG-	3776
Qy	1322	AAAAACCCCATCTGTTAAAGTTATAAGGTGACTTAAACAAATAACATAATAGATGGGGTTTC	1381
Db	3777	-----CTGCCTTTATTGAAAGACAATAATAGACATGTCGGGGTGTA	3816
Qy	1382	TTTTTAATATTATGTGTCCTAATAGTAGACATTTATTCAGATGAAAAATCAAGGGTTTTAGT	1441
Db	3817	TAAACCAATTTTAAAGCGGGATATCGCGCCATCATCCAGATAGTGATACAAGTTTGCTGAA	3876
Qy	1442	GGACAACAAAAAGTGGAAAAGTGAGACCATGTGCTTAGGAGACAGTTATTAAATAGC	1501
Db	3877	AGATATGTCCAGA	3890
Qy	1502	TGAATAGAAGCGGTCTCTCCAAATATCTTTATTAGAAAAGCAAAATCTAAAAATTAATCTG	1561
Db	3891	-----CCAGTCTTTCTCATGAAGACATTAAGCCTTTTGCTTCAA	3930
Qy	1562	AAAAGGAATGAGNAATAGTGNATGGACCAATAATAATGACTAGAGAAGAAAGAAATGAAGA	1621
Db	3931	GGCGCCGCTTGCACCATACGTGTCAGGGAACTTTAGGGAAAGACTGTCAACCATGGAAGA	3990
Qy	1622	TTGTTTCATGAATTAAGGAACGAATATTGATAAATATGGGATGATGTTAAGGCTATTG	1681
Db	3991	GGTTTTAAGCCATTGGGGCGGATTAGAGAAAA-----CATGAATGTTTCATCGTAG	4043
Qy	1682	GTGTTTATGCTCTCTTGGTCGTGAGACTGATGGGCCCTATTCCGATATTGAGATGATGT	1741
Db	4044	AAGTGCAGCGGGTATTTCTGTGCCATTGGAGAGGACTATTTGGTCAGTCATGTCAATA	4103
Qy	1742	GTGTGATGTCAACAGAGGAAGCAGAGTTACGCCATGAATGGCAACCGGTGAGTGGAAAG	1801
Db	4104	-----AAGCGTTGCAGCTTCCCATGATTATTGTGCGCGTCTCGCCTTGGAAAC	4153
Qy	1802	TGGAAGTGAATTTTGATAGCGAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAG	1861
Db	4154	ATTAAATCATACCTTTTAACTGTCAAAATATGCAGAAAGCATTGGGGCTTCCAATCGCCGGA	4213
Qy	1862	ATTGGCGCTTACACATGGTCAATTTTCTCTATTGTTCCGATTATGATTTCAAGTGGAT	1921
Db	4214	ATTATCATCAATGGAATCAGTGACTCTCCCTGATGAAGATGAATAAACAATCTGAGATG	4273
Qy	1922	ACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAAGCCCAAAACGTTCCACGATGCGA	1981
Db	4274	ATTGAGCGCTTATGCGGTGTGCCGATTTTAGGGGTTACGCCAAAGCTTGCCAAACGTG---	4330
Qy	1982	TTTGTGCCCTTATCGTATAGAGAGCTGTTTGAATATGACAGGCAATGGCGTAATATTCGTG	2041
Db	4331	-----ACGAAAGAAACCGGTTCTACATATGTTGTTAAAGACATA-----	4367
Qy	2042	TGCAAGACCGACAACATTTCTACCATCTTTGACTGTACAGTAGCAATGGCAGGTGCCA	2101
Db	4368	-----TCATCTATCATTTACTGATGTAATCAAGTGGGGGTATGA	4405
Qy	2102	TGTTGATGTGTCGATCATCGCATCTGTTTATACGACAGCGCTTCGGTCTTAACTGAAG	2161

Db	4406	GAATGAATCAATGATGGAACTCGCAGACGGGGTCTGGCTGGAGCAGAAAGTGACTGACG	4465
Qy	2162	CAGTTAAAGCAATCAGATCTTCCTTCAGGTTATGACCATCTGTGCCAGTTCGTATGTGCTG	2221
Db	4466	AAGAGGCGCTTTCAAATTAATCATTTGCTCTGATGAAGATAATTTGCTATTAATGCAACGGG	4525
Qy	2222	GTCAACTTTTCGACTCTGAGAAACCTTCTGGAACTCGTAGAAGATTTCTGGAAATGGGATTC	2281
Db	4526	CTTTTCACATCAGAAAAACATCTTTTACGGAAAAAAGTAAAGCTCAATATGATATATGAATG	4585
Qy	2282	AGGAGTGGACAGAAACGACACGGATATATAGTGGATGTCAAAACGCATACCATTTTGTAA	2341
Db	4586	CGAAATCGGGCTCTGCCGGAAACTCGCGCTATTGTTTCA-----GTCTGCGAT	4637
Qy	2342	TTCGAAAGCCCGATTGAGTCTTACCGGATGGTGAAATGAAGAAAACGCTGCTTGAAGCGC	2401
Db	4638	TTCGAAAGCCCGATTGAGTCTTACCGGATGGTGAAATGAAGAAAACGCTGCTTGAAGCGC	4697
Qy	2402	GAAGCGGGCGCACGATCTGAATATCGGCACATATTGTATCGTGGCAAGCGCAGAGGTCC	2461
Db	4698	GAAGCGGGCGCACGATCTGAATATCGGCACATATTGTATCGTGGCAAGCGCAGAGGTCC	4757
Qy	2462	GTCTAAACAGAGAACTGATCAGGTCTGTAGATCGGTTTTCAGAAATTAAGAGACGTATGG	2521
Db	4758	GTCTAAACAGAGAACTGATCAGGTCTGTAGATCGGTTTTCAGAAATTAAGAGACGTATGG	4817
Qy	2522	ACTGAAGATTGTGCATGTCTTGGACTGTGTGAAGCCAGACGACGCGAAGCGGCTCAAAGA	2581
Db	4818	ACTGAAGATTGTGCATGTCTTGGACTGTGTGAAGCCAGACGACGCGAAGCGGCTCAAAGA	4877
Qy	2582	TGCAGGAGTAGACCGCTATTAATCAATAATTGGAATACGTCAAGAGAAAACATTCAAAACAT	2641
Db	4878	TGCAGGAGTAGACCGCTATTAATCAATAATTGGAATACGTCAAGAGAAAACATTCAAAACAT	4937
Qy	2642	CACAACTCTCACATACATACGATGACAGGTCAATACGTTTGAATCGCAAAAGAAATCGGG	2701
Db	4938	CACAACTCTCACATACATACGATGACAGGTCAATACGTTTGAATCGCAAAAGAAATCGGG	4997
Qy	2702	GCTGTCTCCGTGTTTTCAGCGCCCATTTATCGGGATGAAGGAGACGAAAACAGGATGTATTGA	2761
Db	4998	GCTGTCTCCGTGTTTTCAGCGCCCATTTATCGGGATGAAGGAGACGAAAACAGGATGTATTGA	5057
Qy	2762	CATCGCCAAAAGCTTTGAAGGCTCTTGAACGGGATTCATTCCTGTGTGAATTTTTTGTGATGC	2821
Db	5058	CATCGCCAAAAGCTTTGAAGGCTCTTGAACGGGATTCATTCCTGTGTGAATTTTTTGTGATGC	5117
Qy	2822	AATTGATGGACCGCGTTTGAAGCGGTCAACGAATTAACCCGCTGTATTGTTTAAAAGT	2881
Db	5118	AATTGATGGACCGCGTTTGAAGCGGTCAACGAATTAACCCGCTGTATTGTTTAAAAGT	5177
Qy	2882	GCTGGCGCTGTTCGGTTTATCAATCCATCAAAAGAAATTCGCAATTTCCGGAGGAAGAGA	2941
Db	5178	GCTGGCGCTGTTCGGTTTATCAATCCATCAAAAGAAATTCGCAATTTCCGGAGGAAGAGA	5237
Qy	2942	GGTCAATCTCCGCACATGTGACGCCATPAGGGCTTTACCGCGAAAACCTCAATTTTGTGCGG	3001
Db	5238	GGTCAATCTCCGCACATGTGACGCCATPAGGGCTTTACCGCGAAAACCTCAATTTTGTGCGG	5297
Qy	3002	AGACTACTTAACAACTCCCGGGCAAGAGGAGACGGAGGATCATATAAATGCTGAGTGATTT	3061
Db	5298	AGACTACTTAACAACTCCCGGGCAAGAGGAGACGGAGGATCATATAAATGCTGAGTGATTT	5357
Qy	3062	AGGCTTTTGAAGTTGAATCAGTTCGAAGAAATGAAGGCTAGTTTAAAGTCGGAAGAGCTGAAA	3121
Db	5358	AGGCTTTTGAAGTTGAATCAGTTCGAAGAAATGAAGGCTAGTTTAAAGTCGGAAGAGCTGAAA	5417
Qy	3122	GAATCAATAAAGCAATCGGTATGATGTCGAAT	3154
Db	5418	GAATCAATAAAGCAATCGGTATGATGTCGAAT	5450

BSU51868
LOCUS 10153 bp DNA linear BCT 25-OCT-1996
DEFINITION Bacillus subtilis biotin biosynthetic operon genes, complete and partial cds.
ACCESSION U51868
VERSION U51868.1 GI:1277024
KEYWORDS
SOURCE
ORGANISM Bacillus subtilis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 10000)
AUTHORS Bower,S., Perkins,J.B., Yocum,R.R., Howitt,C.L., Rahaim,P. and Pero,J.
TITLE Cloning, sequencing, and characterization of the Bacillus subtilis biotin biosynthetic operon
JOURNAL J. Bacteriol. 178 (14), 4122-4130 (1996)
PUBMED 8763940
AUTHORS Yocum,R.
TITLE Direct Submision
JOURNAL Submitted (20-FEB-1996) Yocum R., Omnigene Bioproducts, Inc., 763 D Concord Ave., Cambridge, MA, USA, 02139-9002
FEATURES
source
1..10153
/organism="Bacillus subtilis"
/mol_type="genomic DNA"
/db_xref="taxon:1423"
/map="262 degrees"
/clones="pBIO100, pBIO350, pBIO201"
2119..2898
/gene="bioW"
2119..2898
/gene="bioW"
/function="biotin pathway"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="pimelyl CoA synthetase"
/protein_id="AAB17457.1"
/db_xref="GI:1277025"
/translation="MMQETFYSVRMSMNGSHEDGGKHISGGERLIPHEMKHTYN ALLBKSHSGKDFPMQIQEPEVHSIKTIQPLPVHTNEVSCPEBGQKLRLLEKE GVSVDTEKAYEQIPEVSDVRGAVLFDIHTGKMDQTKEGVRSVMDWPDANPEKWA LHSVPAHSRIKEALAKASKVSRHPAVVAELCWSDDPDYLTGYVAGKKMGVQRITAKK EYTEGRCRVFDGSDNVNTYIHDLKQPLIIEWEDHDS"
2888..4234
/gene="bioA"
2888..4234
/gene="bioA"
/function="biotin pathway"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DAPA aminotransferase"
/protein_id="AAB17458.1"
/db_xref="GI:1277026"
/translation="MTWDLIEKSKHLWLPFTQMKNVDENPLIESTGTGKVKDINK BYFDGFWLNVGHRKLELDDAIKKQLGKIAHSTLLGMNTNVPATQLAETLIDISPK KLTFRVSDSAGEAMEIALKMAFYKNIKGKEQKFTAMKNGYHGTQITGAVSYGSIE LFHHVYGLPMEYKAPTPYVYRSESGDPDECRCQCLRELAAQLLEEHEEHTAALSIES MVQASGMI VMPGCVLAGVRELCTTYDVLMI VDBVATGFGRTGKMFACENHVOPDLM AAGKGTGGYLP IAVTATEDIYKAFYDDYENLATFFGHSHYTCNQIGCAVALENLAL PSENIYEQVAEKSKLHFLQDLHALPHVGDIRQLGFMCGAELVRSKTEKYPYPADR RIGYVSLKMLGMLNTRPLGVDVIAFLPPLASTAEISEMVAIMKQAIHEVTSLED"
4224..5393
/gene="bioP"
4224..5393
/gene="bioP"
/function="biotin pathway"
/standard_name="KAPA synthase"
/codon_start=1
/evidence=experimental
/transl_table=11

/product="L-alanine - pimelyl CoA ligase"
/protein_id="AAB17459.1"
/db_xref="GI:1277027"
/translation="MKIDSWNLERLDRMKEAGVHRNLRSMGDGAPVPERNIDGENQTWV SSNNYGLASDRRLIDAAQTALQOFGTSGSRLTTGNSVMHEKLEKXIASFUKLTEAA LLFSSGYLANVGLSLPEKEDVILSDLNHSMIDGCRILSKADTVVYRHDMNDLEN KLTSEYQRRFIVTDGVSMDGTIAPLDQIISLAKRYHAFVVVDDAHATGLVDSGQ GTSEYQVCPIDIVIGTLLSKVAGEGGPAAGSAVFIDFLNHARTFIQTALPPASCAA AHEAFNIEASREKRQLLFSYISMIRTSKMGVYVKGDTHTPIIPVIGDAHKTVLFA EKLOGKGIYAPAIRPPTVAPGESIRITITSDHSMGDIIDLHLQTFHSIGELH11"
5390..6085
/gene="bioD"
5390..6085
/gene="bioD"
/function="biotin pathway"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="deathiobiotin synthetase"
/protein_id="AAB17460.1"
/db_xref="GI:1277028"
/translation="MRGFFVIGTDEVGKTVISSGLAALLKONNRHVGVYKPFLLSGIS RHPDSDTSLKDMSQTSLSHEDITPFAKAPAPYVAGLEKGTVMEEVLSHWGR1 REKHECFIVGAGGISVPLGEDYLVSHVIALQLPMIIVARPLGTINHTLTVKYAE SMGLPIAGIITINGISDSDPEDEKTNPEMERLCGPILGVTPKLANVTKETVLHMVKD HINLSLLMNQGV"
6088..7095
/gene="bioB"
6088..7095
/gene="bioB"
/function="biotin pathway"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="biotin synthase"
/protein_id="AAB17461.1"
/db_xref="GI:1277029"
/translation="MNQWMLADRVLAGAEVDEALSILHCPDEDILLMHGAFH1R KHFYKKVYKLANIMNAKSLGPCNCGYCSAISKAPIESYRVNWKETLLEGAKRAHD LNI GTYCVASGRPSNRVDQVVDVAQVEIKETYGKICACGLGLKPEQAKRLKADGV ARKHLNLTDSQRHNSNITTSHTYDDRVNTVEIAKESGLSPCSGAILGKMTKQVDI DI KSLKALDADSIPVNFILHAIDGTFLEGVNELNPLYCLKVLALFRFINPSKEIRISGGR EVNRLTLQPLGLYAANSIFVGDVLTITAGQEBTEDEHKMLSDLGFEVESVEEMKASLSAK S"
7164..8351
/gene="bioI"
7164..8351
/gene="bioI"
/function="pimelic acid synthesis"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="cytochrome P450-like enzyme"
/protein_id="AAB17462.1"
/db_xref="GI:1277030"
/translation="MTIASSTASSEFLKNPVSFYDTLRAVHPIYKGSFLKYPGVVVTG YETAILLKDKARKVITPPLPESSTKYQDLSDHVNQWMLFQNDPHRLRLTLASGFTP RTTESQPIYIETVHLLDDQKQKQMEVISEDFAFLPASFVIANIIGVPEDRQOLKE WAASLQITDITFTRKALTEGNTMAVQAMAYFKELIQKRKHPOODMISMLLKGREK KLTEEEAASCTILLATAGHETVTNLTISNVLCLIQHPQLKLENDPDLIGTAVEECL RYSPQMTARVASEDIDICGVTIROGEQVYLLLGAA NRDPISFTNPDPFDITRSPN HLSFGHHVYCLGSSLARLEAQIAINTLLQRMFSLNLADFEWRYRPLFGFRALBELPV TFS"
8428..9189
/gene="orf2"
8428..9189
/gene="orf2"
/function="unknown"
/codon_start=1
/transl_table=11
/product="unknown, probably a redox enzyme"
/protein_id="AAB17463.1"
/db_xref="GI:1277031"

Db 5834 ATTAATCATACCTTTTAACTGTCAAAATATGCAGAAAGCATGGGCTCTTCCAATCGCGGA 5893
Qy 1862 ATTGGCCGCTTACACATGGTCAATTTTCTCTATTTTTCGCCATTTATGATTCAGGTGGAT 1921
Db 5894 ATTATCATCAATGGAATCAGTGACTCTCTCTGATGAAGATGAAAAACCAATCCTGAGATG 5953
Qy 1922 ACTTAGAGAAAGTGTATCAAACTGTAATCGGTAGAGCCCAACGTTCCACAGTGGGA 1981
Db 5954 ATTGAGCGCTTATGCGGTGTCGCGATTTTAGGGTTACGCCAAGACTTGCACACGTG--- 6010
Qy 1982 TTTGTGCCCTTATCGTAGAGAGCTGTTTGAATATGACGGAATAATGGCGTAAATTCGTG 2041
Db 6011 -----ACGAAGAAGAACGGTTCTACATATGTTAAAGACCAT- 6047
Qy 2042 TGCAGAGGACGACAAACATTTTACACATCTTGAAGCTGTATACAGTAGCAATGCGAGTGCCA 2101
Db 6048 -----TCAATCTATCATTTACTGATGNAATCAAGTGGGGGTATGA 6085
Qy 2102 TGTGATTGGTCTGCATCATCGCATCTGTATATACGACGAGCGCTTCGGTCTTTAACTGAAG 2161
Db 6086 GAATGAATCAATGATGGAATCTGCAGACCGGTGCTGGCTGGAGCAGAAGTGACTGACG 6145
Qy 2162 CAGTTAAGCAATCAGATCTTCTTTCAGGTTATGACCAATCTGTGCCAGTTTCGTAATGTCG 2221
Db 6146 AAGAGGCGCTTTCRAATATTACATTTGCTGCTGATGAAGATATTTTGTCTAATATGACGCGG 6205
Qy 2222 GTCAACTTTCGACTCTGAGAAACTTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTC 2281
Db 6206 CTTTTCACATCAGAAAACACTTTTACGGAAAAAAGTAAAGCTCAATATGATTATGAATG 6265
Qy 2282 AGGAGTGACAGAACGACACCGATATATAGTGGATGTGTCAAAAGCGCATACCATTTTGA 2341
Db 6266 CGAATCGGCGCTCTGCCGGAANAACCTCGGCTATTTGTACA-----GTCTGCGAT 6317
Qy 2342 TTCGAAAGCGCGATTGAGTCTTACCGGATGGTGAATAAGGAAACGCTGCTTGAAGCGGC 2401
Db 6318 TTCGAAAGCGCGATTGAGTCTTACCGGATGGTGAATAAGGAAACGCTGCTTGAAGCGGC 6377
Qy 2402 GAAGCGGCGCAGCATCTGAATATCGGCACATATGTATCTGTGCGAAGCGCAGAGTCC 2461
Db 6378 GAAGCGGCGCAGCATCTGAATATCGGCACATATGTATCTGTGCGAAGCGCAGAGTCC 6437
Qy 2462 GTCTAACAGAGAGTGGATCAGTCTAGATGCGGTTCAGGAAATTAAGAGACGTATGG 2521
Db 6438 GTCTAACAGAGAGTGGATCAGTCTAGATGCGGTTCAGGAAATTAAGAGACGTATGG 6497
Qy 2522 ACTGAAGATTTGTGCATGCTTTGGACTGTTGAAGCCAGAGCGGAAGCGGCTCAAGA 2581
Db 6498 ACTGAAGATTTGTGCATGCTTTGGACTGTTGAAGCCAGAGCGGAAGCGGCTCAAGA 6557
Qy 2582 TGCAGGATAGCCGCTATATCATATTTGAATACGTACAGAGAAACCAATTCAAACAT 2641
Db 6558 TGCAGGATAGCCGCTATATCATATTTGAATACGTACAGAGAAACCAATTCAAACAT 6617
Qy 2642 CACAACCTCACATACATACGATGACAGAGTCAATACGTTCAATTCGAAATCGCAAGAAATCGG 2701
Db 6618 CACAACCTCACATACATACGATGACAGAGTCAATACGTTCAATTCGCAAGAAATCGG 6677
Qy 2702 GCTGTCTCCGTGTTTCAGCGCCATTTATCGGATGAAGGAGACGAAACAGGATGTCATGA 2761
Db 6678 GCTGTCTCCGTGTTTCAGCGCCATTTATCGGATGAAGGAGACGAAACAGGATGTCATGA 6737
Qy 2762 CATGCCAAAAGCTTGAAGGCTCTTGACCGGATCCATTCCTGTGAATTTTTTGCATGC 2821
Db 6738 CATGCCAAAAGCTTGAAGGCTCTTGACCGGATCCATTCCTGTGAATTTTTTGCATGC 6797
Qy 2822 AATTGATGCCAGCGGTTAGAGGCGTCAACGAATTTAAACCCGCTGTATTTGTTAAAGT 2881
Db 6798 AATTGATGCCAGCGGTTAGAGGCGTCAACGAATTTAAACCCGCTGTATTTGTTAAAGT 6857
Qy 2882 GCTGCGCTGTTTCGTTTTTATCAATCCATCAAAAGAAATTCGCAATTTCCGAGGAAGAGA 2941

Db 6858 GCTGGCGCTGTTCCGTTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAAGAGA 6917
Qy 2942 GGTCAATCTCGCACATTGCGACCATTAGGCTTTTACGCGCAAACTCCATTTTGTGCGG 3001
Db 6918 GGTCAATCTCGCACATTGCGACCATTAGGCTTTTACGCGCAAACTCCATTTTGTGCGG 6977
Qy 3002 AGACTACTTAACTACTCCCGGCAAGAGGAGACGAGGATCATAAATTCGTGAGTGATTT 3061
Db 6978 AGACTACTTAACTACTCCCGGCAAGAGGAGGAGGAGGATCATAAATTCGTGAGTGATTT 7037
Qy 3062 AGGCTTTTGAATGATCAGTCGAGAAATGAAGGCTAGTTTAAAGTGGAAAAAGCTGAAA 3121
Db 7038 AGGCTTTTGAATGATCAGTCGAGAAATGAAGGCTAGTTTAAAGTGGAAAAAGCTGAAA 7097
Qy 3122 GAATCAATAAAGCAATCGGTATGATGTCGAAT 3154
Db 7098 GAATCAATAAAGCAATCGGTATGATGTCGAT 7130

RESULT 7
BSUB0016/c
LOCUS
DEFINITION
BSUB0016 199922 bp DNA linear BCT 18-APR-2005
Bacillus subtilis complete genome (section 16 of 21): from 3013458 to 3213379.
ACCESSION
VERSION
Z99119 AL009126
Z99119.2 GI:32468809

SOURCE
ORGANISM
REFERENCE
AUTHORS
Bacillus subtilis subsp. subtilis str. 168
Bacillus subtilis subsp. subtilis str. 168
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
1 (bases 1 to 199922)

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriass, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerston, I.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emmerger, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J., Grandi, G., Guilbepp, G., Guy, B.J., Haga, K., Haeck, J., Harwood, C.R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F., Itoya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaere-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moesti, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portetelle, D., Porwollik, S., Prescott, A.M., Prescann, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serrero, P., Shin, B.S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Takakoshi, A., Tanaka, T., Terpetra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenberg, M., Vannier, F., Vassartoli, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenecker, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zumstein, E., Yoshikawa, H. and Danchin, A.

TITLE
The complete genome sequence of the gram-positive bacterium
Bacillus subtilis
JOURNAL
PUBMED
NATURE 390 (6657), 249-256 (1997)
9384377
2 (bases 1 to 199922)
REFERENCE
Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
Direct Submission
Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur, Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr, adanchin@pasteur.fr
Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48

COMMENT

On Jul 7, 2003 this sequence version replaced gi:2635411.
This entry contains data from release R16.1 of the Subtilist
database. Further data on gene annotation and detailed information
about changes from previous releases can be found at
<http://genolist.pasteur.fr/Subtilist/>.

FEATURES

source

location/Qualifiers

1. .199922

/organism="Bacillus subtilis subsp. subtilis str. 168"

/mol_type="genomic DNA"

/strain="168"

/db_xref="taxon:224308"

/complement(96..608)

/gene="moaB"

/locus tag="BSU29460"

/complement(96..608)

/gene="moaB"

/locus tag="BSU29460"

/codon_start=1

/transl_table=11

/product="molybdopterin precursor biosynthesis protein B"

/protein_id="CAB14924.1"

/db_xref="GI:2635430"

/db_xref="GOA:O34457"

/db_xref="UniProt/Swiss-Prot:O34457"

/translation="MSVQEHKQEPDIVRCKVITSDTRTDETDKSGKLMISPLEEAG

HIIAIIYVDEKDALSVLAGMDEQTDVALLNGGTGIADRDVTEATPFSKEL

PGFGEIPRLMSYTEDIGSAIIMSRATAGVIQHTAVFTSGSGGAVKAMNKLIPELA

HVVRIRKDK"

/complement(693..1880)

/gene="ackA"

/locus tag="BSU29470"

/complement(693..1880)

/gene="ackA"

/locus tag="BSU29470"

/EC number="2.7.2.1"

/function="conversion of acetyl-CoA to acetate"

/codon_start=1

/transl_table=11

/product="acetate kinase"

/protein_id="CAB14925.1"

/db_xref="GI:2635431"

/db_xref="GOA:P37877"

/db_xref="InterPro:IPR000890"

/db_xref="InterPro:IPR004372"

/db_xref="Subtilist:BG10813"

/translation="MSKIIAIINAGSSSLKQLFEMPSETVLTGKLVIRIGIADSVFTI

SVNGKTEVTDIPDAVAVKMLNKLTEFGIILKDLNEIDGHRVHGGEKFSDSL

LTQETKEIEDISELAPLNPANLVGIIKAFKVLPNVAVPDTAFHQTMPQSVLY

SLPYEYKGIKYGFGTSHKVTYTERAAELGRLPKDLRLISCHLNGASIAVVG

SKSDTSMTGTPLAGVAGTRSGNIDPALIPIYMEKGTQDADEVLTNKKSGLLGIS

GFSSDLDIYEATKEGNEARALEVFASRIHKYISYAARMSGVDAIIFTAGIGENS

VEVREVLRLGLEFVGWYDPAIENVGEEAFISYPHSPVKVMIPTDEVMIAADVVR

LAK"

/complement(2228..3217)

/gene="ytxK"

/locus tag="BSU29480"

/complement(2228..3217)

/gene="ytxK"

/locus tag="BSU29480"

/function="unknown"

/note="alternate gene name: ythI

similar to unknown proteins"

/codon_start=1

/transl_table=11

/protein_id="CAB14926.1"

/db_xref="GI:2635432"

/db_xref="GOA:P37876"

/db_xref="InterPro:IPR000051"

/db_xref="InterPro:IPR002296"

gene

CDS

/db_xref="Subtilist:BG10812"

/db_xref="UniProt/Swiss-Prot:P37876"

/translation="MOKDHVAVYELLNEAAIMIKNELQISYIEALAEAGEMYFLEKT

DOLKPADOKTKQALKEKAEFTYEHVWRKAFOLAVLKGMDKISHPNRQMTPTDI

GLFISIVNKPMDAKKELTILDPALGTGNLLFTVNLQSKTANSFGIEIDDLVLLKIA

YAQNLKKELELPHQDSLEPLFIDPVDTVICDLPVGYIYENDEGAFAELAKDEGHSF

AHLFIQEVSVKHTPGGYLFFMIPNHLFESSQSGKLQFFKDKVHINALLQPKSIFK

DEAHAKSILVLRQGGENTKAPGQILLANLPSFSNQKAMLDMAQDFDEWFKKEK"

/complement(3278..3781)

/gene="tpx"

/locus tag="BSU29490"

/complement(3278..3781)

/gene="tpx"

/locus tag="BSU29490"

/note="alternate gene name: ytgI"

/codon_start=1

/transl_table=11

/product="thiol peroxidase"

/protein_id="CAB14927.1"

/db_xref="GI:2635433"

/db_xref="GOA:P80864"

/db_xref="Subtilist:BG13855"

/db_xref="UniProt/Swiss-Prot:P80864"

/translation="MAEITPKGGPVLVGGVKGQDAPDFTVLTNSLEKSLADMKG

KVTIISVIPSIDTGVCDQATRFNEEAAGLDVNVVYTSADLPFAQARWCGANGIDKV

ETLSDHRDMSFGSAFGVYIKELRLARSFVLDENGKVYAYVSEATNHPNVEKPIE

AAKALVK"

/complement(3891..4346)

/gene="ytfJ"

/locus tag="BSU29500"

/complement(3891..4346)

/gene="ytfJ"

/locus tag="unknown"

/function="unknown"

/note="similar to unknown proteins"

/codon_start=1

/transl_table=11

/protein_id="CAB14928.1"

/db_xref="GI:2635434"

/db_xref="Subtilist:BG13849"

/db_xref="UniProt/TREMBL:O34806"

/translation="MADHPIQGLMKMTAMENLKEMIDVNTIIGDPVETPDGSGVILTQSK

VGFGAAGSGFGGPAEKSEDETRQKLPFGGSGGVSITPIAFLIVGSGTIRM

LHLDENTHLTEKILDAAPQTLERIQOMFKKNKNKQSQGQNNQNNNY"

/complement(4360..5040)

/gene="ytfI"

/locus tag="BSU29510"

/complement(4360..5040)

/gene="ytfI"

/locus tag="unknown"

/function="unknown"

/note="similar to unknown proteins"

/codon_start=1

/transl_table=11

/protein_id="CAB14929.1"

/db_xref="GI:2635435"

/db_xref="Subtilist:BG13848"

/db_xref="UniProt/TREMBL:O34917"

/translation="WYVLTAILLIGVILLRLMKIHVALEYLHADNDKLTALITLAL

FGILIRIKTIIPAKVKNKDGTVDIKORTESMKKSGKKKITQNDVQGSINKLEMLH

QVTRLRKISASFLAHLITKLEWTVIGIQDAAVTGLTGAVNGKGGIAAMLYDLHD

FVNKPYVEIPSPQVPVSKTHFCIFFRFGHAMLAFAKFKVYKVGRELSVLKQNPARM

TSKNDSSV"

/complement(5115..5609)

/gene="yteJ"

/locus tag="BSU29520"

/complement(5115..5609)

/gene="yteJ"

/locus tag="BSU29520"

/function="unknown"

/note="similar to unknown proteins"

/codon_start=1

/transl_table=11

gene

CDS

Qy 2102 TGTGTAATGGTCTGCATCATCGCATCTCTTTATACGACGAGCGCTTCGGTCTTAACTGAAG 2161
Dy |||||
77083 GAATGAATCAATGATGGAATCTGCAGACCGGGTCTGGCTGGAGCAGAACTGACTGACG 77024
Qy 2162 CAGTTAAGCAATCAGATCTTCCTTCAGGTTATGACCATCTGTGCGAGTTCTGTAATGCTG 2221
Dy |||||
77023 AAGAGGCGCTTCAATATATACATTTGTCTGATGAAGATATTTTGTCTATTATGACACGGG 76964
Qy 2222 GTCAACTTTCGACTCTGAGAAACTTCTGGAATCGCTAGAGAAATTTCTGGAATCGGATTC 2281
Dy |||||
76963 CTTTTTCACATCAGAAACACTTTTACGGAATAAAGTAAAGCTCAATATGATGATGATG 76904
Qy 2282 AGGAGTGGACAGAACGACCGATATATAGTGGATGTGTCAAAACGACATACCATTTTGA 2341
Dy |||||
76903 CGAAATCGGGCTCTGCCCGGAAACTCGCGCTATTGTTTACA-----GTCTCGCAT 76852
Qy 2342 TTCCAAAGCCCGATGAGTCTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGCGC 2401
Dy |||||
76851 TTCAAAGCCCGGATGAGTCTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGCGC 76792
Qy 2402 GAAGCGGCGCAGCATCTGAATATCGGCACATATTGTATCGTGCACAGCGGCAGAGTCC 2461
Dy |||||
76791 GAAGCGGCGCAGCATCTGAATATCGGCACATATTGTATCGTGCACAGCGGCAGAGTCC 76732
Qy 2462 GTCTAACAGAGAAGTGGATCAGTCTGATAGTCCGTTTCAGGAATTTAAAGACAGCTATGG 2521
Dy |||||
76731 GTCTAACAGAGAAGTGGATCAGTCTGATAGTCCGTTTCAGGAATTTAAAGACAGCTATGG 76672
Qy 2522 ACTGAAGATTGTGATGCTTGTGACTTGTGAAGCCAGAGCAGCGAAGCGGCTCAAAGA 2581
Dy |||||
76671 ACTGAAGATTGTGATGCTTGTGACTTGTGAAGCCAGAGCAGCGAAGCGGCTCAAAGA 76612
Qy 2582 TGCAGGAGTACACGCTATATATCAATTTTCAATACGTTCACAGAGAAACCAATTCAAACAT 2641
Dy |||||
76611 TGCAGGAGTACACGCTATATATCAATTTTCAATACGTTCACAGAGAAACCAATTCAAACAT 76552
Qy 2642 CACAACCTCACATACATACATGATGACAGTCAATACGGTTGAAATTCGCAAAAGAAATCGGG 2701
Dy |||||
76551 CACAACCTCACATACATGATGACAGTCAATACGGTTGAAATTCGCAAAAGAAATCGGG 76492
Qy 2702 GCTGTCTCCGTGTTCCAGCGCATATTCGGATGAAGAGACGAAACAGGATGTCATGA 2761
Dy |||||
76491 GCTGTCTCCGTGTTCCAGCGCATATTCGGATGAAGAGACGAAACAGGATGTCATGA 76432
Qy 2762 CATGCCAAAGCTTGAAGGCTCTTGACGCGATTCCATTCCTGTGAATTTTTCGATGC 2821
Dy |||||
76431 CATGCCAAAGCTTGAAGGCTCTTGACGCGATTCCATTCCTGTGAATTTTTCGATGC 76372
Qy 2822 AATTGATGGCACGCGTTAGAGCGGTCAACGAATTTAAACCCGCTGTATTGTTTAAAGT 2881
Dy |||||
76371 AATTGATGGCACGCGTTAGAGCGGTCAACGAATTTAAACCCGCTGTATTGTTTAAAGT 76312
Qy 2882 GCTGGCGTGTTCGTTTTATCAATTCATCAAAAGAAATTCGATTTCCGAGAGAGAGA 2941
Dy |||||
76311 GCTGGCGTGTTCGTTTTATCAATTCATCAAAAGAAATTCGATTTCCGAGAGAGAGA 76252
Qy 2942 GGTCAATCTCCGACATTCGAGCCATTAGGCTTTTACGCGCAACTCCATTTTGTTCGG 3001
Dy |||||
76251 GGTCAATCTCCGACATTCGAGCCATTAGGCTTTTACGCGCAACTCCATTTTGTTCGG 76192
Qy 3002 AGACTACTTAAACAATCGCGGCAAGAGGACGAGGATCATATAATTCGTGAGTGATTT 3061
Dy |||||
76191 AGACTACTTAAACAATCGCGGCAAGAGGACGAGGATCATATAATTCGTGAGTGATTT 76132
Qy 3062 AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTGGAAAGCTGAA 3121
Dy |||||
76131 AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTGGAAAGCTGAA 76072
Qy 3122 GAATCAATAAAGCAATCGGTATGATGTCGAT 3154
Dy |||||
76071 GAATCAATAAAGCAATCGGTATGATGTCGAT 76039

RESULT 8
AF008220
LOCUS 22060 bp DNA linear BCT 04-FEB-1998
DEFINITION Bacillus subtilis rrnB-dnaB genomic region.
ACCESSION AF008220
VERSION AF008220.1 GI:2293135
KEYWORDS
SOURCE
ORGANISM Bacillus subtilis
Bacillus subtilis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 7430)
AUTHORS Green,C.J., Stewart,G.C., Hollis,M.A., Vold,B.S. and Bott,K.F.
TITLE Nucleotide sequence of the Bacillus subtilis ribosomal RNA operon, rrnB
Gene 37 (1-3), 261-266 (1985)
REFERENCE 2 (bases 153210 to 153762)
AUTHORS Connors,M.J., Mason,J.M. and Setlow,P.
TITLE Cloning and nucleotide sequencing of genes for three small, acid-soluble proteins from Bacillus subtilis spores
J. Bacteriol. 166 (2), 417-425 (1986)
REFERENCE 3 (bases 213161 to 218473)
AUTHORS Ogasawara,N., Moriya,S., Mazza,P.G. and Yoshikawa,H.
TITLE Nucleotide sequence and organization of dnaB gene and neighbouring genes on the Bacillus subtilis chromosome
Nucleic Acids Res. 14 (24), 9989-9999 (1986)
REFERENCE 4 (bases 200404 to 201481)
AUTHORS Saki,T., Yoshikawa,H., Takahashi,H. and Saito,H.
TITLE Cloning and nucleotide sequence of phoP, the regulatory gene for alkaline phosphatase and phosphodiesterase in Bacillus subtilis
J. Bacteriol. 169 (7), 2913-2916 (1987)
REFERENCE 5 (bases 201234 to 203212)
AUTHORS Saki,T., Yoshikawa,H., Takahashi,H. and Saito,H.
TITLE Nucleotide sequence of the Bacillus subtilis phoR gene
J. Bacteriol. 170 (12), 5935-5938 (1988)
REFERENCE 6 (bases 142232 to 144147)
AUTHORS Grundy,F.J. and Henkin,T.M.
TITLE Cloning and analysis of the Bacillus subtilis rpsD gene, encoding ribosomal protein S4
J. Bacteriol. 172 (11), 6372-6379 (1990)
REFERENCE 7 (bases 133624 to 134990)
AUTHORS Henkin,T.M., Grundy,F.J., Nicholson,W.L. and Chambliss,G.H.
TITLE Catabolite repression of alpha-amylase gene expression in Bacillus subtilis involves a trans-acting gene product homologous to the Escherichia coli lacI and galP repressors
Mol. Microbiol. 5 (3), 575-584 (1991)
REFERENCE 8 (bases 140810 to 142610)
AUTHORS Henkin,T.M., Glass,B.L. and Grundy,F.J.
TITLE Analysis of the Bacillus subtilis tyrS gene: conservation of a regulatory sequence in multiple trRNA synthetase genes
J. Bacteriol. 174 (4), 1299-1306 (1992)
REFERENCE 9 (bases 217570 to 220060)
AUTHORS Putzer,H., Gendron,N. and Grunberg-Manago,M.
TITLE Co-ordinate expression of the two threonyl-tRNA synthetase genes in Bacillus subtilis: control by transcriptional antitermination involving a conserved regulatory sequence
EMBO J. 11 (8), 3117-3127 (1992)
REFERENCE 10 (bases 134990 to 141290)
AUTHORS Grundy,F.J., Waters,D.A., Takova,T.Y. and Henkin,T.M.
TITLE Identification of genes involved in utilization of acetate and acetoin in Bacillus subtilis
Mol. Microbiol. 10 (2), 259-271 (1993)
REFERENCE 11 (bases 162129 to 164080)

Db	87857	TTCCGAAGCGCCGATTGAGTCTTTACCGGATGTGTAATAGGAACCGCTGCTTTGAAGGGCG	87916
Qy	2402	GAAGCGGGCGCAGATCTCAATATCGGCACATATTGTATCGTGGCAAGCGGCAGAGGTCC	2461
Db	87917	GAAGCGGGCGCAGATCTCAATATCGGCACATATTGTATCGTGGCAAGCGGCAGAGGTCC	87976
Qy	2462	GTCTAACAGAGAAGTGGATCAGGTCTGTAGATCGGTTTCAGGAAATTTAAAGAGACGTATGG	2521
Db	87977	GTCTAACAGAGAAGTGGATCAGGTCTGTAGATCGGTTTCAGGAAATTTAAAGAGACGTATGG	88036
Qy	2522	ACTGAAGATTTGTGCATGCTTTGGACATGTTTGAAGCAGAGCAGCGAAGCGGCTCAAGA	2581
Db	88037	ACTGAAGATTTGTGCATGCTTTGGACATGTTTGAAGCAGAGCAGCGAAGCGGCTCAAGA	88096
Qy	2582	TGCAGGAGTAGACCGCTATAATCATATTTGAATACGTACAGAGAAACCATTTCAAACAT	2641
Db	88097	TGCAGGAGTAGACCGCTATAATCATATTTGAATACGTACAGAGAAACCATTTCAAACAT	88156
Qy	2642	CACAACTTCACATACATACGATGACAGAGTCAATACGGTTGAAATTCGCAAAAGATCGGG	2701
Db	88157	CACAACTTCACATACATACGATGACAGAGTCAATACGGTTGAAATTCGCAAAAGATCGGG	88216
Qy	2702	GCTGTCTCCGTGTTTCAGGCGCCATTATTCGGGATGAAGGAGACGAAACAGGATGTCATTGA	2761
Db	88217	GCTGTCTCCGTGTTTCAGGCGCCATTATTCGGGATGAAGGAGACGAAACAGGATGTCATTGA	88276
Qy	2762	CATGCCCAAAAGCTTGAAGGCTCTTGACCGGATTCCTTCCTGTGAATTTTTTGCATGC	2821
Db	88277	CATGCCCAAAAGCTTGAAGGCTCTTGACCGGATTCCTTCCTGTGAATTTTTTGCATGC	88336
Qy	2822	AATTGATGGCAGCGCTTAGAAGCGCTCAACGAAATTTAAACCCGCTGTATTGTTTAAAGT	2881
Db	88337	AATTGATGGCAGCGCTTAGAAGCGCTCAACGAAATTTAAACCCGCTGTATTGTTTAAAGT	88396
Qy	2882	GCTGGCGCTGTTCCGTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAAGAGA	2941
Db	88397	GCTGGCGCTGTTCCGTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAAGAGA	88456
Qy	2942	GGTCAATCTCCGACATTCGAGCCATTAGGCGTTTACGCCGCAAACTCCATTTTGTGCG	3001
Db	88457	GGTCAATCTCCGACATTCGAGCCATTAGGCGTTTACGCCGCAAACTCCATTTTGTGCG	88516
Qy	3002	AGACTACTTTAACTGCGGGCAAGAGGAGACGAGGATCAATAAATGCTGAGTGATTT	3061
Db	88517	AGACTACTTTAACTGCGGGCAAGAGGAGACGAGGATCAATAAATGCTGAGTGATTT	88576
Qy	3062	AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTCGAAAGCTGAAA	3121
Db	88577	AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTCGAAAGCTGAAA	88636
Qy	3122	GAATCAATAAAGCAATCGGTATGATGCGAAT	3154
Db	88637	GAATCAATAAAGCAATCGGTATGATGCGAAT	88669
RESULT 9			
LOCUS	AB088066	6739 bp	DNA linear BCT 17-JUL-2004
DEFINITION	Bacillus subtilis biotin operon (bioW, bioA, bioF, bioD, bioB, bioI), complete cds.		
ACCESSION	AB088066		
VERSION	AB088066.1 GI:21734667		
KEYWORDS	Bacillus subtilis		
SOURCE	Bacillus subtilis		
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
REFERENCE	1		
AUTHORS	Sasaki, M., Kawamura, F. and Kurusu, Y.		
TITLE	Genetic analysis of an incomplete bio operon in a biotin		
JOURNAL	auxotrophic strain of Bacillus subtilis natto OK2		
PUBMED	Biosci. Biotechnol. Biochem. 68 (3), 739-742 (2004) 15056910		

REFERENCE	2	(bases 1 to 6739)
AUTHORS	Kurusu, Y., Sasaki, M. and Kotanagi, T.	
TITLE	Direct Submission	
JOURNAL	Submitted (10-JUL-2002) Yasurou Kurusu, School of Agriculture, Ibaraki University, Laboratory of Molecular Microbiology, 3-21-1, Chuo, Ami, Inashiki, Ibaraki 300-0393, Japan	
FEATURES	(E-mail: krsy@ipc.ibaraki.ac.jp, Tel: 81-298-88-8646)	
source	Location/Qualifiers	
	1..6739	/organism="Bacillus subtilis"
		/mol_type="genomic DNA"
		/isolate="OK2"
		/db_xref="taxon:1423"
		/note="synonym:Bacillus natto"
gene	561..1238	/gene="bioW"
CDS	561..1238	/gene="bioW"
		/codon_start=1
		/product="6-carboxyhexanoate-CoA ligase"
		/protein_id="BAC03239.1"
		/db_xref="GI:21734668"
		/translations="MMQBEFFYSVVRMRASMGSHEDGKHISGGERLIPHEMKHTVN ALLEKGLSHSRGKPFQMIQIEEVHESIKTIQPLPYVTNEVSCPEGQKLARLLEKE LVSRDVEIKAYEQIPEWSDVRGAVLFDIHTGKMDQTKBKGVVRVSRMDWPDANFEKWA LHSHPAHSYRIKEALALASKVSRHPAAIAELCWSDDPDYITGTVAGKMGVQYRITAMK EYGTESG"
gene	1330..2676	/gene="bioA"
CDS	1330..2676	/gene="bioA"
		/gene="bioA"
		/codon_start=1
		/product="adenosylmethionine-8-amino-7-oxononanoate aminotransferase"
		/protein_id="BAC03240.1"
		/db_xref="GI:21734669"
		/translations="MTHDLLEKSKHLWLPFTOMKYDENPLIIESGTGKIVQDINGK EYDGRSSVWLVGHKKELDNAIKQLCKIAHSTLLGNTVNPATOLAETLIDISPK KLFTRVYSDGSAEMALAKMAFOYKNIGKPEKQFIAMKNGYHGTIGAVSVGSIE LFHHVYPLMFSEYKAPIPYVYRSESDGDECCQLELEQLLKEHEEIAALUSIES IVGASGMIYMPGEYLAGVRELCTTYDVLMIYDEVTGFGTGMFAHENVQFDLM AAGKGTGGYLPVATFATEDIYKAFYDVENLKTFFHSGTYGNOGCAVALNADL PEGENVEOVAEKSKLHLFLODLHALPHVGDTRQLGFMCGAELVRSKETKEPYPADR RICYKSLKMRLEGLTRPLGDVIAFLPPLASTAEELTEMVAIMKQAIHEVTSLEN"
gene	2666..3781	/gene="bioF"
CDS	2666..3781	/gene="bioF"
		/gene="bioF"
		/codon_start=1
		/product="8-amino-7-oxononanoate synthase"
		/protein_id="BAC03241.1"
		/db_xref="GI:21734670"
		/translations="MKIDSWLNDRLDIAKEAGVHRNLRSMNGAPVPERNIDGENQTVW SSNNYLGASDRDLIDAAQTALQOFGTSGSGSRLLTGNVHKEKLEKIASFPKTEAA LFPSSGLANVGLVLSLPEKEDVILSDLANHASIIDGRLSKADTVVYRHTDMNDLEN KLNTRYQRRFIVTGVFSMDGTIVPLDQIISLAKRYHAFVYVDDAHATCVLGDGSR GTSEYGVYDDIVIGTSLKAVGTGGGAAGSAVFDFLNLHARTFIQTALPPASCAA AHEAFNRTSLKNGYVVGKDHTPIIPVIVIGDAKTVIFAEKLGQGGKGIYAPAIRPPTV APGESRIRITITSDHSMGDIDDLKLTFFHSIGKELHI"
gene	3778..4473	/gene="bioD"
CDS	3778..4473	/gene="bioD"
		/codon_start=1
		/product="dethiobiotin synthetase"
		/protein_id="BAC03242.1"
		/db_xref="GI:21734671"
		/translations="MRGFFVTGTDTEVGKTVISSGLAALLKEHNRDGVYKPFSLGIS RHHPDSDTSLKDMSQTSLSHEDITPPAFKAPLAPYVAGKLGKTVTNEEVLHWGRI

Qy	1562	AAAAGGAA	TGAGAAT	TAGTGAAT	TGAGCA	CAATA	TAATA	TGACTAG	GAGAA	GAAAGAA	TGAAGA	1621
Db	3999	GGCCCGCT	TGCA	CCATAC	GTG	TG	CAGGAA	AACTT	GAGGAA	GAAGACT	GTGCAC	4058
Qy	1622	TTGTTCAT	GAAATTA	AGAA	CGAATAT	TGGA	TAAAT	TGGGAT	TGAT	TGGGAT	TGATGT	1681
Db	4059	GGTTTTAG	CCATT	TGGGG	CGGAT	TAG	AAAAA	-----	-----	-----	CATGA	4097
Qy	1682	GTGTTTAT	TGGCTCT	TTTGGT	CGTCAG	ACTG	TATGG	CCCTAT	TTCGGAT	TATTTG	AGATGATGT	1741
Db	4098	ATGCTTC	ATCGTAG	AAGGT	CGGG	CGGTAT	TTCTGT	GCCTAT	TTCGGAG	GAGCA	TATTTTGGT	4157
Qy	1742	GTGTCTAT	GTCA	CAGAGAA	GACGAG	TTCAG	CCATGA	ATGGA	CAACCG	TGTGAG	TGGAAGG	1801
Db	4158	CAGTCAT	GTCTATA	AAAGCG	TTCAG	CTTCC	CATGAT	TATTTG	TGCGCG	GTCTC	CACCTTGG	4217
Qy	1802	TGGAAGT	GAAATTTT	TGATAG	CGAAGAG	ATTCTA	CTAGAT	TATG	CATCTC	CAGT	TGGAATCAG	1861
Db	4218	AACCAT	TAACTCAT	ACCTTTT	TAACTGT	CAAA	TATG	CAGAA	AGCAT	-----	-----	4266
Qy	1862	ATTGGCCG	CTTAC	CATGGT	CAATTTT	TCTAT	TTTTC	CGCAT	TATTA	TGATTC	AGGTGGAT	1921
Db	4263	----GGGG	CTTCA	ATTG	CCGGA	TATC	-----	-----	-----	ATCA	ATGGAATCAGT	4306
Qy	1922	ACTTAGA	GAAGTGT	ATCAAA	CTCTAA	ATCG	GTAGA	AGCCAA	ACGTTT	CCAGAT	CGGA	1981
Db	4307	CTCCTG	ATGA	ATGA	AAAAAAC	CAATC	CTGAG	ATGA	TGAG	CGCTTAT	GCGTGTCCGA	4366
Qy	1982	TTTGTGCC	CTTAT	CGTAG	AAGAG	CTGTTT	TGNA	TATG	CAGGCA	ATGCG	GTAAATTCGTG	2041
Db	4367	TTTTAG	GGGTTAC	CCCAAA	-----	-----	GC	TGCCAA	CGTTAC	GAAGAA	AAACGGTTCTA	4416
Qy	2042	TGCAAG	ACCGACA	CAATTTCT	ACCAT	CTCTG	ACTGTAC	AGTAG	CAATG	GCAGGT	CGCA	2101
Db	4417	CATATG	TATA	AAAGCA	TATCA	TCTAT	CA	TACTG	---ATGA	TCAAT	TGGGGGTATGA	4473
Qy	2102	TGTTGAT	TGGTCTG	CA	TCATCG	CA	CTGTTAT	CAC	GAGCG	CGCTTC	CGGTCTTAACTGAAG	2161
Db	4474	GAATGA	ATCA	TGAT	GSNA	CTCG	CAAA	CGGGT	CTGG	TGGAG	CAGAGTGA	4533
Qy	2162	CAGTTA	AGCAAT	CAGAT	CTTCC	TTCAG	GTATG	ACCA	ATCTGT	GCCAG	TTCGTATATGTCG	2221
Db	4534	AAGAG	CGCTTTT	TCAAT	TATTA	CTTGT	CTGAT	GAAG	ATATTTT	TGCTAT	TAAATGCA	4593
Qy	2222	GTCAACT	TTTCCG	ACTCTG	GAGAA	CTTCT	GGAAT	CGCTAG	AAGATTT	CTGGA	ATGGGA	2281
Db	4594	CTTTTCA	TATCAG	AAAAACA	CTTTTAT	TG	AAAA	AAAGTAA	AGCTCA	ATATG	ATATGA	4653
Qy	2282	AGGAGT	GGACAG	AACG	ACAGATAT	TAGT	GGATGT	GTCA	AAACG	CAATAC	CAATTTTGA	2341
Db	4654	CGAAT	CCGGCT	CTG	CCCG	AAAACT	TG	CGGCTAT	TGTTCA	-----	-----GTCTG	4705
Qy	2342	TTCAAA	AGCGCG	AT	TGAGT	CTTAC	CGGAT	TGGTGA	ATAAG	AAACG	CTGCTTGA	2401
Db	4706	TTCAAA	AGCGCG	AT	TGAGT	CTTAT	CGAT	TGGT	GAATAAG	AAACG	CTGCTTGA	4765
Qy	2402	GAAG	CGGCGCAC	CAATCT	CGAATAT	CGG	CACATAT	TGTAT	CGTGC	CAAGCG	CAGAGTCC	2461
Db	4766	GAAG	CGGCGCAC	CAATCT	CGAATAT	CGG	CACATAT	TGTAT	CGTGC	CAAGCG	CAGAGTCC	4825
Qy	2462	GTCTAA	CAGAGA	AGTGG	ATCAG	GTCTG	TAGAT	TGCG	GGTTT	CAGGAA	TATTAAG	2521
Db	4826	GTCTAA	CAGAGA	AGTGG	ATCAG	GTCTG	TAGAT	TGCG	GGTTT	CAGGAA	TATTAAG	4885
Qy	2522	ACTGAA	GATTTG	TGCAT	GTCTT	TGGA	CTTGTGA	AGCCAG	ACGCG	GAAGCG	GTCAAAGA	2581
Db	4886	ACTGAA	GATTTG	TGCAT	GTCTT	TGGA	CTTGTGA	AGCCAG	ACGCG	GAAGCG	GTCAAAGA	4945
Qy	2582	TGCAG	GAGT	TAG	CCGCTAT	ATAT	TCAAT	TATG	CAATCG	TACAG	GAAACCA	2641
Db	4946	TGCCG	GAGT	TAG	CCGCTAT	ATAT	TCAAT	TATG	CAATCG	TACAG	GAAACCA	5005
Qy	2642	CACAA	CTCTCA	ATACAT	ACGAT	GACAG	AGTCA	ATAC	CGGTT	GAAAT	TCGAAAA	2701

D	b		S	006	CACGACCTCACAATCATATCGATGACAGAGTCAATACGGTTGAAATCCGAAGAAGTAATCGGG	5065
Q	y		7	02	GCTGTCTCCGTGTTTCAGGCGCCATTATCGGGATGAAGGAGACGAAAAAGGATGTCATTGA	2761
D	b		5	06	GCTGTCTCCGTGTTTCAGGCGCCATTATCGGGATTAAGGAGACGAAAACAGATGTCATTGA	5125
Q	y		2	72	CATCGCCAAAAGCTTTGAAGGCTCTTGACGCCGATTCATTCTGTGTGAATTTTTTTGCATGC	2821
D	b		5	12	CATCGCCAAAAGCTTTGAAGGCTTTGACGCCGATTCATTCTGTGTGAATTTTTTTGCATGC	5185
Q	y		2	82	AATTGATGGCACCGCGTTTAGAAGCGTCAACGAATTAAACCCGCTGTATTGTTTAAAAGT	2881
D	b		5	18	AATTGATGGCACCGCGTTTAGAAGCGTCAACGAATTAAACCCGCTGTATTGTTTAAAAGT	5245
Q	y		2	88	GTCGGCGCTGTTCCGTTTTTATCAATCCCATCAAAGAAATTCGCAATTTCCGGAGGAAGAGA	2941
D	b		5	24	GCTGGCGCTGTTCCGTTTTTATCAATCCCATCAAAGAAATTCGCAATTTCCGGAGGAAGAGA	5305
Q	y		2	94	GGTCAACTCTCCGCACATTCGAGCCATTAGGGCTTTACGCCGCAAACTCCAATTTTTGTGCG	3001
D	b		5	30	GGTCAACTCTCCGCTCAATTCGAGCCATTAGGGCTTTACGCCGCAAACTCCAATTTTTGTGCG	5365
Q	y		3	02	AGACTACTTAAACAATCTCCGGCAGAGGAGACGGAGGATCATATAATGCTEAGTGATTT	3061
D	b		5	36	AGACTACTTAAACAATCTCCGGCAGAGGAGACGGAGGATCATATAATGCTEAGTGATTT	5425
Q	y		3	06	AGGCTTTGAAAGTTGAAATCAGTCGAAGAAATGAAGGCTAGTTTAAGTCGGAAGAGCTCAAA	3121
D	b		5	42	AGGCTTTGAAAGTTGAAATCAGTCGAAGAAATGAAGGCTAGTTTAAGTCGGAAGAGCTCAAA	5485
Q	y		3	12	GAATCAATAAAAGCAATCGGTATGATGTCGAATTC	3156
D	b		5	48	GAATCAATAAAAGCAATCGGTATGATGTCGAATTC	5520
RESULT 10						
AX825002						
LOCUS AX825002 3560 bp DNA linear PAT 11-DEC-2003						
DEFINITION Sequence 1 from Patent WO03072785.						
ACCESSION AX825002						
VERSION AX825002.1 GI:39750777						
KEYWORDS synthetic construct						
SOURCE other sequences; artificial sequences.						
ORGANISM						
REFERENCE 1 Hohmann,H.P., Mouncey,N.J., Sauer,U. and Zamboni,N.						
AUTHORS Fermentation process						
TITLE Patent: WO 03072785-A 1 04-SBP-2003;						
JOURNAL Roche Vitamins AG (CH)						
FEATURES						
source .locus= "synthetic construct"						
/mol_type= "unassigned DNA"						
misc_feature 1..3560 /db_xref= "taxon:32630"						
ORIGIN /note= "plasmid pNMR21"						
Query Match 38.6%; Score 1219; DB 6; Length 3560;						
Best Local Similarity 99.5%; Pred. No. 5.7e-254;						
Matches 1235; Conservative 0; Mismatches 0; Indels 6; Gaps 1;						
Q	y		1	02	AGCTTGGGCGACAGGTCGAGATCGGAATGAGTTTATAAATAAAAAAGCACCTGAAA	1161
D	b		1	46	AGCTTGGGCGACAGGTCGAGATCGGAATGAGTTTATAAATAAAAAAGCACCTGAAA	1525
Q	y		1	62	AGGTGTCCTTTTTTGATGGTTTTGAACCTGTCCTTCTTATCTTGATACATATAGAAATA	1221
D	b		1	26	AGGTGTCCTTTTTTGATGGTTTTGAACCTGTCCTTCTTATCTTGATACATATAGAAATA	1585
Q	y		1	22	ACGTCAATTTTTATTTTTTATTTAGTTTCTCGTGAAGGTCGTTGAAGGTTTGGTATGTATGT	1281

Db 1586 ACATC-----ATTATTTATTTAGTGTCTGAAAGGTGCGTTGAAAGTGTGGTATGTATGT 1639
Qy 1282 GTTTTAAAGTATTGAAACCCCTTAAATTTGGTGCACAGAAACCCCATCTGTTAAAGT 1341
Db 1640 GTTTTAAAGTATTGAAACCCCTTAAATTTGGTGCACAGAAACCCCATCTGTTAAAGT 1699
Qy 1342 TATAAGTGACTAAACAAATAAATAAGTAGATGGGGTTTCTTTTAATAATTATGTGCTCTA 1401
Db 1700 TATAAGTGACTAAACAAATAAATAAGTAGATGGGGTTTCTTTTAATAATTATGTGCTCTA 1759
Qy 1402 ATAGTAGCAATTTATCAGATGAAATCAAGGGTTTATGTGCAAGCAAAAAAGTGGAA 1461
Db 1760 ATAGTAGCAATTTATCAGATGAAATCAAGGGTTTATGTGCAAGCAAAAAAGTGGAA 1819
Qy 1462 AAGTGAGACCATGTCTTAGAAGACGAGTTTAAATAGCTGAATAGAAAGGGTCTCTC 1521
Db 1820 AAGTGAGACCATGTCTTAGAAGACGAGTTTAAATAGCTGAATAGAAAGGGTCTCTC 1879
Qy 1522 CAATATTTCTTATTTAGAAAGCAAAATCTAAATTTATCTGAAAGGGAATGAGAAATAGT 1581
Db 1880 CAATATTTCTTATTTAGAAAGCAAAATCTAAATTTATCTGAAAGGGAATGAGAAATAGT 1939
Qy 1582 AATGACCAATTAATAGCACTAGAGAGAAAGAAATGAAGTTGTTCAATGAATTAAGGAA 1641
Db 1940 AATGACCAATTAATAGCACTAGAGAGAAAGAAATGAAGTTGTTCAATGAATTAAGGAA 1999
Qy 1642 CGAATATTTGGATAAATATGGGATGATGTTAAGCTATTGGTGTATGGCTCTCTTGGT 1701
Db 2000 CGAATATTTGGATAAATATGGGATGATGTTAAGCTATTGGTGTATGGCTCTCTTGGT 2059
Qy 1702 CGTCAGACTGATGGGCCCCATTTCGGATATTGAGATGATGTGTGTCATGTCACAGAGGAA 1761
Db 2060 CGTCAGACTGATGGGCCCCATTTCGGATATTGAGATGATGTGTGTCATGTCACAGAGGAA 2119
Qy 1762 GCAGAGTTCAGCCATGAATGAGACAAACCGTGAGTGGAAAGTGGAAATTTGATGAGC 1821
Db 2120 GCAGAGTTCAGCCATGAATGAGACAAACCGTGAGTGGAAAGTGGAAATTTGATGAGC 2179
Qy 1822 GAACAGATTTCTACTAGATATTGATCTCAGGTGGAATCAGATTCGCGCTTACACATGGT 1881
Db 2180 GAACAGATTTCTACTAGATATTGATCTCAGGTGGAATCAGATTCGCGCTTACACATGGT 2239
Qy 1882 CAATTTTTCTCTATTTTGGCGAATTTATGATTCAGGTGGATCTTATGAGAAAGTGTATCAA 1941
Db 2240 CAATTTTTCTCTATTTTGGCGAATTTATGATTCAGGTGGATCTTATGAGAAAGTGTATCAA 2299
Qy 1942 ACTGCTAAATCGGTAGAGCCCAAACGGTTCCAGATGCGAATTTGTCGCCCTTATCGTAGAA 2001
Db 2300 ACTGCTAAATCGGTAGAGCCCAAACGGTTCCAGATGCGAATTTGTCGCCCTTATCGTAGAA 2359
Qy 2002 GAGCTGTTTGAATATGAGGCAAAATGCGTAATATTGCTGTGCAAGGACCGACAAATTT 2061
Db 2360 GAGCTGTTTGAATATGAGGCAAAATGCGTAATATTGCTGTGCAAGGACCGACAAATTT 2419
Qy 2062 CTACCATCTTTGACTGTACAGTAGCAATGCGAGTGCATGTTGATTTGTTGCTGCATCAT 2121
Db 2420 CTACCATCTTTGACTGTACAGTAGCAATGCGAGTGCATGTTGATTTGTTGCTGCATCAT 2479
Qy 2122 CGCATCTGTTATACAGAGCGCTTTCGGTCTTAATCTGAAGCAGTTTAAGCAATCAGATCTT 2181
Db 2480 CGCATCTGTTATACAGAGCGCTTTCGGTCTTAATCTGAAGCAGTTTAAGCAATCAGATCTT 2539
Qy 2182 CCTTCAGTTATGACCATCTGTGCGCAGTTGCTAATGCTGTGTCACCTTCCGACTCTGAG 2241
Db 2540 CCTTCAGTTATGACCATCTGTGCGCAGTTGCTAATGCTGTGTCACCTTCCGACTCTGAG 2599
Qy 2242 AAACCTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGAC 2301
Db 2600 AAACCTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGAC 2659
Qy 2302 GGATATATAGTGGATGTGCAAAACCGCATACCATTTTGAAT 2342

Db 2660 GGATATATAGTGGATGTGCAAAACCGCATACCATTTTGAAT 2700
RESULT 11
LOCUS CQ894464
DEFINITION Sequence 1 from Patent WO2004090138. linear
ACCESSION CQ894464
VERSION CQ894464.1 GI:55467386
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barg, H. and Jahn, D.
TITLE Method for the production of vitamin B12
JOURNAL Patent: WO 2004090138-A 1 21-OCT-2004;
BASF Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
source 1..7350
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
promoter 9..18
/note="xylA-Promotor"
misc_feature 27..953
/note="cbiX B. megaterium"
rep_origin 1123..1173
/note="oriB.coli (pBR322)"
misc_feature 1906..2786
/note="amp-Resistenzgen (counterclockwise)"
rep_origin 3611..3686
/note="oriB.megaterium (pE194ts)"
rep_origin 4008..4615
/note="repF (counterclockwise)"
misc_feature 6399..7169
/note="neo-Resistenzgen"
ORIGIN
Query Match 38.08; Score 1200.4; DB 6; Length 7350;
Best Local Similarity 99.08; Pred. No. 5.8e-250;
Matches 1233; Conservative 0; Mismatches 1; Indels 12; Gaps 2;
Qy 1102 AGCTTGGGACGAGGTGCGAGATCAGGGAATGAGTTTATAAAATAAAAAACACCTCAAA 1161
Db 5925 AGCTTGGCTGCGAGTGCAGATCAGGGAATGAGTTTATAAAATAAAAAACACCTCAAA 5984
Qy 1162 AGGTGCTTTTTTTTGTAGTGGTTTGAACCTGTTCTTCTTATCTTTGATACATATAGAAATA 1221
Db 5985 AGGTGCTTTTTTTTGTAGTGGTTTGAACCTGTTCTTCTTATCTTTGATACATATAGAAATA 6044
Qy 1222 ACCTCATTTTTTATTTTATTTTATTTAGTTGCTCAAGAGTGGTGGTGGTATGTATGT 1281
Db 6045 ACCTC-----ATTATTTATTTTATTTAGTTGCTCAAGAGTGGTGGTGGTATGTATGT 6098
Qy 1282 GTTTTAAAGTATTGAAACCCCTTAAATTTGGTGCACAGAAACCCCATCTGTTAAAGT 1341
Db 6099 GTTTTAAAGTATTGAAACCCCTTAAATTTGGTGCACAGAAACCCCATCTGTTAAAGT 6158
Qy 1342 TATAAGTGACTAAACAAATAAATAAGTAGATGGGGTTTCTTTTAATAATTATGTGCTCTA 1401
Db 6159 TATAAGTGACTAAACAAATAAATAAGTAGATGGGGTTTCTTTTAATAATTATGTGCTCTA 6218
Qy 1402 ATAGTAGCAATTTATCAGATGAAATCAAGGGTTTATGTGGAACAAGCAAAAAAGTGGAA 1461
Db 6219 ATAGTAGCAATTTATCAGATGAAATCAAGGGTTTATGTGGAACAAGCAAAAAAGTGGAA 6278
Qy 1462 AAGTGAGACCATG-----TGCTTAGGAAGACGAGTTTATTAATAGCTGAATTAAGAACGGT 1515
Db 6279 AAGTGAGACCATGAGCTTATGCTTAGGAAGACGAGTTTATTAATAGCTGAATTAAGAACGGT 6338
Qy 1516 GCTCTCCAAATATCTTATTTATTTAGAAAGCAAAATCTAAATTTATCTGAAAAAGGGAATGAGA 1575


```
QY 2064 ACCATCCTTGACTGTACAGGTAGCAATGGCAGGTGCATGTTGATTGGTCTGCATCATCG 2123
Db 5260 ACCATCCTTGACTGTACAGGTAGCAATGGCAGGTGCATGTTGATTGGTCTGCATCATCG 5319
QY 2124 CATCTGTTATACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGATCTTCC 2183
Db 5320 CATCTGTTATACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGATCTTCC 5379
QY 2184 TTCAGGTTATGACCATCTGTGCCAGTTCGTAATGCTCTGGTCAACTTTCGGACTCTGAGAA 2243
Db 5380 TTCAGGTTATGACCATCTGTGCCAGTTCGTAATGCTCTGGTCAACTTTCGGACTCTGAGAA 5439
QY 2244 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGAATTCAGGAGTGGACAGAACGACACGG 2303
Db 5440 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGAATTCAGGAGTGGACAGAACGACACGG 5499
QY 2304 ATATATAGTGGATGTGCAAAACGCATACCATTTTGA 2341
Db 5500 ATATATAGTGGATGTGCAAAACGCATACCATTTTGA 5537

RESULT 13
CS115022/c
LOCUS CS115022 8644 bp DNA linear PAT 08-JUL-2005
DEFINITION Sequence 3 from Patent WO2005056799.
ACCESSION CS115022
VERSION CS115022.1 GI:70663705
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Nielsen, A.K. and Rasmussen, M.D.
TITLE A cell with improved secretion mediated by mrga protein or homologue
JOURNAL Patent: WO 2005056799-A 3 23-JUN-2005;
Novozymes A/S (DK)
FEATURES
    Location/Qualifiers
        source
            1..8644
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Plasmid pDG268neo"

ORIGIN
Query Match 37.8%; Score 1192.8; DB 6; Length 8644;
Best Local Similarity 98.9%; Pred. No. 2.5e-248;
Matches 1224; Conservative 0; Mismatches 7; Indels 7; Gaps 2;

QY 1105 TTGGGCAGCAGGTCGAGATCAGGAATGAGTTTATAAATAAAGCAACCTGCAAAAGG 1164
Db 6664 TGGAGGAGTATGGGGAGATCAGGAATGAGTTTATAAATAAAGCAACCTGCAAAAGG 6605
QY 1165 TGTCTTTTGTGAGTTTGAATCTGTTCTTCTTATCTTGATACATATAGAAATAACG 1224
Db 6604 TGTCTTTTGTGAGTTTGAATCTGTTCTTCTTATCTTGATACATATAGAAATAACG 6545
QY 1225 TCATTTTATTTTATTTTATTTAGTTCTGCTGAAAGGTCGGTGAAGTGTGGTATGATGTT 1284
Db 6544 TC-----ATTTTATTTTATTTAGTTCTGCTGAAAGGTCGGTGAAGTGTGGTATGATGTT 6491
QY 1285 TTAAGTATTGAAACCTTAAATTTGTTTCACAGAAACCCCATCTGTTAAGTTAT 1344
Db 6490 TTAAGTATTGAAACCTTAAATTTGTTTCACAGAAACCCCATCTGTTAAGTTAT 6431
QY 1345 AAGTGACTTAAACAATACTAAATAGATGGGGTTTCTTTTAAATATTATGTCCTTAATA 1404
Db 6430 AAGTGACTTAAACAATACTAAATAGATGGGGTTTCTTTTAAATATTATGTCCTTAATA 6371
QY 1405 GTAGCATTTATTCAGATGAAATAATCAAGGGTTTTAGTGGACAGACAAAAGTGGAAAAG 1464
Db 6370 GTAGCATTTATTCAGATGAAATAATCAAGGGTTTTAGTGGACAGACAAAAGTGGAAAAG 6311
```

```
QY 1465 TGAGACCATG-TGCTTAGGAAGACGAGTTATTATATAGCTGAATAAGAACGGTGTCTTCCA 1523
Db 6310 TGAGACCATGATCTTCTTAGGAAGACGAGTTATTATATAGCTGAATAAGAACGGTGTCTTCCA 6251
QY 1524 AATATTTCTTTATTTAGAAAAGCAATCTTAAATTTATCTGAAAAGGGGAATGAGATAGTAA 1583
Db 6250 AATATTTCTTTATTTAGAAAAGCAATCTTAAATTTATCTGAAAAGGGGAATGAGATAGTAA 6191
QY 1584 TGACCAATAATAATGACTAGAGAAAGAAATGAAGATTTGTTTCATGAAATTAAGGAACG 1643
Db 6190 TGACCAATAATAATGACTAGAGAAAGAAATGAAGATTTGTTTCATGAAATTAAGGAACG 6131
QY 1644 AATATTTGGATTAATATGGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTTGGTCG 1703
Db 6130 AATATTTGGATTAATATGGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTTGGTCG 6071
QY 1704 TCAGACTGATGGGCCCTTATTCGGATTTAGAGATGATGTCATGTCACACAGGAAGC 1763
Db 6070 TCAGACTGATGGGCCCTTATTCGGATTTAGAGATGATGTCATGTCACACAGGAAGC 6011
QY 1764 AGAGTTTCAGCCATGAATGGACAACCGGTGAGTGGAAAGTGGAAATTTTGATAGCGA 1823
Db 6010 AGAGTTTCAGCCATGAATGGACAACCGGTGAGTGGAAAGTGGAAATTTTGATAGCGA 5951
QY 1824 AGAGATTTCTACTAGATTTATGCATCTCAGGTGGAAATCAGATTGGCCGCTTACACATGTCA 1883
Db 5950 AGAGATTTCTACTAGATTTATGCATCTCAGGTGGAAATCAGATTGGCCGCTTACACATGTCA 5891
QY 1884 ATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGATCTAGAGAAAGTGTATCAAC 1943
Db 5890 ATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGATCTAGAGAAAGTGTATCAAC 5831
QY 1944 TGCTAAATCGGTAGAGCCCAAAAGTTCCACAGATGCGATTGTGCCCCTTATCGTAGAGA 2003
Db 5830 TGCTAAATCGGTAGAGCCCAAAAGTTCCACAGATGCGATTGTGCCCCTTATCGTAGAGA 5771
QY 2004 GCTGTTTGAATATCGAGGCAATGGCGTAATATTTCGTGTCAGAGGACCGACAATTTCT 2063
Db 5770 GCTGTTTGAATATCGAGGCAATGGCGTAATATTTCGTGTCAGAGGACCGACAATTTCT 5711
QY 2064 ACCATCTTGAATGATCAGGTAGCAATGGCAGGTGCATGTTGATTGGTCTGCATCATCG 2123
Db 5710 ACCATCTTGAATGATCAGGTAGCAATGGCAGGTGCATGTTGATTGGTCTGCATCATCG 5651
QY 2124 CATCTGTTATACGACGAGCGCTTCGGTCTTAACTGGAAGCAGTTAAGCAATCAGATCTTCC 2183
Db 5650 CATCTGTTATACGACGAGCGCTTCGGTCTTAACTGGAAGCAGTTAAGCAATCAGATCTTCC 5591
QY 2184 TTCAAGGTTATGACCATCTGTGCCAGTTCGTAATGCTCTGGTCAACTTTCGGACTCTGAGAA 2243
Db 5590 TTCAAGGTTATGACCATCTGTGCCAGTTCGTAATGCTCTGGTCAACTTTCGGACTCTGAGAA 5531
QY 2244 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGAATTCAGGAGTGGACAGAACGACACGG 2303
Db 5530 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGAATTCAGGAGTGGACAGAACGACACGG 5471
QY 2304 ATATATAGTGGATGTGCAAAACGCATACCATTTTGA 2341
Db 5470 ATATATAGTGGATGTGCAAAACGCATACCATTTTGA 5433

RESULT 14
CS091382/c
LOCUS CS091382 13129 bp DNA linear PAT 03-JUN-2005
DEFINITION Sequence 8 from Patent WO2005042750.
ACCESSION CS091382
VERSION CS091382.1 GI:66948711
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
```


D	b	2480	AATATGGGAGTATGTTAAGGCATTATGGTGTTATATGGCTCTCTTTGGTCGTACAGTCGATG	2539
Q	y	1715	GGCCCTATTTCGGATATTGAGATGATGTGTGCATGTCAAACAGAGGAAGCAGAGTTTCAGGCC	1774
D	b	2540	GGCCCTATTTCGGATATTGAGATGATGTGTGCATGTCAAACAGAGGAAGCAGAGTTTCAGGCC	2599
Q	y	1775	ATGAATCGACAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGTATAGCGAAGAGATCTTAC	1834
D	b	2600	ATGAATCGACAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGTATAGCGAAGAGATCTTAC	2659
Q	y	1835	TAGATTATGCATCTCACGGTGGAAATCAGATTTGGCGGCTTACACATGGTCAATTTTTCTCTA	1894
D	b	2660	TAGATTATGCATCTCACGGTGGAAATCAGATTTGGCGGCTTACACATGGTCAATTTTTCTCTA	2719
Q	y	1895	TTTTTGGCGATTTATGATTTACAGTGGATCTATTAGAGAAAGTGATCAAACTGCTAAAATCGG	1954
D	b	2720	TTTTTGGCGATTTATGATTTACAGTGGATCTATTAGAGAAAGTGATCAAACTGCTAAAATCGG	2779
Q	y	1955	TAGAAGCCCCAAAACGTTCCAGATCGCATTTGTGCCCCTTATCGTAGAAGAGCTGTTTGAAT	2014
D	b	2780	TAGAAGCCCCAAAACGTTCCAGATCGCATTTGTGCCCCTTATCGTAGAAGAGCTGTTTGAAT	2839
Q	y	2015	ATGCAGAGCAAAATGCGCTAAATATTTCGTGTGCAGAACGACGCAACAACTTCTACCACCTTGA	2074
D	b	2840	ATGCAGAGCAAAATGCGCTAAATATTTCGTGTGCAGAACGACGCAACAACTTCTACCACCTTGA	2899
Q	y	2075	CTGTACAGGTAGCAATGGCAGGTGCCATGTGTGATTTGGTCTGCATCATCGCATCTGTTATA	2134
D	b	2900	CTGTACAGGTAGCAATGGCAGGTGCCATGTGTGATTTGGTCTGCATCATCGCATCTGTTATA	2959
Q	y	2135	CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTCCAGGTTATG	2194
D	b	2960	CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTCCAGGTTATG	3019
Q	y	2195	ACCATCTGTCCAGTTCGTAATGTCTCGTCAAATTTCCGACTCTGAGAACTTCTCGAAT	2254
D	b	3020	ACCATCTGTCCAGTTCGTAATGTCTCGTCAAATTTCCGACTCTGAGAACTTCTCGAAT	3079
Q	y	2255	CGCTAGAGAAATTTCTGGAATGGGAATTCAGGAGTGGACGAAACGACCGGATATATAGTGG	2314
D	b	3080	CGCTAGAGAAATTTCTGGAATGGGAATTCAGGAGTGGACGAAACGACCGGATATATAGTGG	3139
Q	y	2315	ATGTGTCAAAACGCAATACCAATTTTGA	2341
D	b	3140	ATGTGTCAAAACGCAATACCAATTTTGA	3166

Search completed: February 11, 2006, 21:06:23
Job time : 10503 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2006, 17:36:45 ; Search time 1217 Seconds
(without alignments)
17283.295 Million cell updates/sec

Title: US-10-681-086-1
Perfect score: 3156
Sequence: 1 ggatccacagggttagcagc.....atcggtatgatgtcgaaattc 3156

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*
- 14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3156	100.0	3156	6 ABL57129	AbL57129 Bacillus
2	1612.6	51.1	8478	2 AAO81792	AaQ81792 B. subtilis
3	1219	38.6	3560	10 ACF36386	ACF36386 Nucleotid
4	1200.4	38.0	7350	13 ADT25766	Adt25766 Vitamin B
5	1192.8	37.8	8152	14 AEB16616	Aeb16616 Plasmid p
c 6	1192.8	37.8	8644	14 AEB16603	Aeb16603 Plasmid p
c 7	1192.8	37.8	13129	14 AEA06372	Aea06372 Novel bac
8	1081	34.3	1170	13 ADT48504	Adt48504 Bacterial
9	867	27.5	4334	14 AEA43211	Aea43211 M. enegma
c 10	867	27.5	5142	5 AAF29645	Aaf29645 Plasmid p
11	867	27.5	5793	6 AAS17998	Aas17998 Plasmid p
12	867	27.5	5865	12 ADL16337	Adl16337 pBP11 bac
13	867	27.5	5943	6 AAS17997	Aas17997 pBP11 bac
c 14	867	27.5	6661	10 AAD49392	Aad49392 pMO1995 p
15	867	27.5	6694	12 ADL16336	Adl16336 pBP1 back
16	867	27.5	7311	14 AEA06371	Aea06371 Novel bac
17	867	27.5	7336	11 AAN91856	Aan91856 Plasmid p
18	867	27.5	8119	2 AAT33477	Aat33477 Subtilisi
19	867	27.5	8198	12 ADL16342	Adl16342 pBP103 ex

c 20	867	27.5	8808	2 AAT32227	Aat32227 Plasmid p
c 21	867	27.5	9144	2 AAT18951	Aat18951 Plasmid p
c 22	867	27.5	9286	12 ADL16345	Adl16345 pBP105 ex
c 23	867	27.5	10140	2 AAT32232	Aat32232 Plasmid p
c 24	867	27.5	10929	6 ABK98591	Abk98591 pEP25 vec
c 25	867	27.5	10929	9 ACD13842	Acd13842 Plasmid p
c 26	865.4	27.4	5302	3 AAA98012	Aaa98012 Expressio
c 27	865.4	27.4	5767	3 AAA98014	Aaa98014 Expressio
c 28	865.4	27.4	7334	1 AAN80664	Aan80664 Sequence
29	853	27.0	5609	6 AAD23867	Aad23867 pCFC717 D
30	853	27.0	5609	6 AAD23868	Aad23868 pEB200 DN
31	853	27.0	7026	6 AAD23871	Aad23871 pEB303 pl
32	853	27.0	7344	6 AAD23869	Aad23869 pEB301 ch
33	796.8	25.2	934	2 AAT18374	Aat18374 CrtB gene
34	796.8	25.2	2476	2 AAT18377	Aat18377 Plasmid p
35	768.8	24.4	1365	10 ACD02293	ACD02293 Polymorph
36	762.4	24.2	1200	2 AAV02474	Aav02474 Ant(4')-I
37	762.4	24.2	1200	2 AAV02453	Aav02453 Ant(4')-I
38	762.4	24.2	1200	4 AAC81670	Aac81670 E. coli/S
c 39	762.4	24.2	4775	13 ADU49640	Adu49640 Nucleotid
40	759	24.0	759	4 AAF83650	Aaf83650 S. aureus
41	758	24.0	7456	2 AAQ10686	Aaq10686 Plasmid p
42	701	22.2	6837	6 AAD29902	Aad29902 Plasmid p
c 43	700.2	22.2	8084	14 ADW21121	Adw21121 B. amylol
44	695	22.0	6169	2 AAT39284	Aat39284 Mobilisab
45	521.2	16.5	1133	6 ABK73702	Abk73702 Bacillus

ALIGNMENTS

RESULT 1

ABL57129
ID ABL57129 standard; DNA; 3156 BP.

XX ABL57129;

XX 05-AUG-2002 (first entry)

XX Bacillus subtilis bioFDB gene cassette.

XX Biotin; bioFDB; auxotrophy; riboflavin; fermentation; gene; ds.

XX Bacillus subtilis.

XX CH2353772-A1.

XX 08-FEB-2002.

XX 03-AUG-2001; 2001CA-02353772.

XX 08-AUG-2000; 2000US-00633927.

XX (HOFF) ROCHE VITAMINS AG.

XX Stebbins JW, Schlieker HW, Hohmann H, Mouncey NJ;

XX WPI; 2002-424224/62.

XX Producing a target fermentation product, e.g., riboflavin, comprises providing a fermentation medium and a recombinantly-produced microorganism comprising a mutation that causes auxotrophic growth of the microorganism.

XX Claim 19; Page 27-28/1; 42pp; English.

XX The present sequence is the bioFDB gene cassette of Bacillus subtilis. The present invention provides a process for producing a target fermentation product. The process uses an auxotrophic host microorganism that over-produces the fermentation product. The host microorganism is grown in a culture medium supplied in excess with the substrates required for production of the fermentation product and in growth-limiting amounts of a substrate complementing the auxotrophy. The host cell may be


```
Qy 1981 ATTTGTGCCCTTATCGTAGAGAGCTGTTTGAATATGCAAGCAAAATGAGGCTAGCTTAAAGTGCAGAAAGCTGAA 2040
Db 1981 ATTTGTGCCCTTATCGTAGAGAGAGCTGTTTGAATATGCAAGCAAAATGAGGCTAGCTTAAAGTGCAGAAAGCTGAA 2040
Qy 2041 GTGCAAGGACCGACAACTTCTACCACTCTTGACTGTACAGGTAGCAAGTGCAGGTGCC 2100
Db 2041 GTGCAAGGACCGACAACTTCTACCACTCTTGACTGTACAGGTAGCAAGTGCAGGTGCC 2100
Qy 2101 ATGTTGATGCTCTGCAATCATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACCTGAA 2160
Db 2101 ATGTTGATGCTCTGCAATCATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACCTGAA 2160
Qy 2161 GCAGTTAAGCAATCAGATCTTCTTCAGGTTATGACCACTCTGTGCCAGTTTCGTAATGCTCT 2220
Db 2161 GCAGTTAAGCAATCAGATCTTCTTCAGGTTATGACCACTCTGTGCCAGTTTCGTAATGCTCT 2220
Qy 2221 GGTCAACTTTCGGACTCTGAGAACTTCTGGAATCGTAGAGAAATTTCTGGAATGGGATT 2280
Db 2221 GGTCAACTTTCGGACTCTGAGAACTTCTGGAATCGTAGAGAAATTTCTGGAATGGGATT 2280
Qy 2281 CAGGAGTGGACAGAAACGACGAGATATATAGTGATGTGTCAAAACGACATACCAATTTTGA 2340
Db 2281 CAGGAGTGGACAGAAACGACGAGATATATAGTGATGTGTCAAAACGACATACCAATTTTGA 2340
Qy 2341 ATTGAAAGCGCCGATTGAGTCTTACCGGATGTTGAATAGGAAACGCTGCTTGAAGCGC 2400
Db 2341 ATTGAAAGCGCCGATTGAGTCTTACCGGATGTTGAATAGGAAACGCTGCTTGAAGCGC 2400
Qy 2401 CGAAGCGGGCGACGATCTGNAATTCGGACATATTCGTATCGTGGCAAGCGCGAGAGTC 2460
Db 2401 CGAAGCGGGCGACGATCTGNAATTCGGACATATTCGTATCGTGGCAAGCGCGAGAGTC 2460
Qy 2461 CGTCTACAGAGAGTGGATCGAGTCTGTAGATCGGTTTCAGGAAATTAAGAGACGATG 2520
Db 2461 CGTCTACAGAGAGTGGATCGAGTCTGTAGATCGGTTTCAGGAAATTAAGAGACGATG 2520
Qy 2521 GACTGAAGATTGTCATGCTTGGACTGTTGAGCCAGGACGAGCGGCAAGCGGCTCAAAG 2580
Db 2521 GACTGAAGATTGTCATGCTTGGACTGTTGAGCCAGGACGAGCGGCAAGCGGCTCAAAG 2580
Qy 2581 ATGCAGAGTAGACCGCTTAATCATATTTGAATAGTCAACAGAGAAACCATTCAAACA 2640
Db 2581 ATGCAGAGTAGACCGCTTAATCATATTTGAATAGTCAACAGAGAAACCATTCAAACA 2640
Qy 2641 TCACAACCTCACATACATACGATCAGAGTCAATACGTTGAAATCGCAAAAGAAATCGG 2700
Db 2641 TCACAACCTCACATACATACGATCAGAGTCAATACGTTGAAATCGCAAAAGAAATCGG 2700
Qy 2701 GGCTGTCTCGTGTTCAGGCGCCATTTATCGGGATGAAAGGAGACGAAACAGGATGTCATTG 2760
Db 2701 GGCTGTCTCGTGTTCAGGCGCCATTTATCGGGATGAAAGGAGACGAAACAGGATGTCATTG 2760
Qy 2761 ACATCGCAAAAGCTTGAAGGCTTTCAGCGGATTCATTCGTGAAATTTTTCGATG 2820
Db 2761 ACATCGCAAAAGCTTGAAGGCTTTCAGCGGATTCATTCGTGAAATTTTTCGATG 2820
Qy 2821 CAATTGATGCGACCGCTTGAAGCGCTCAACGAATTAACCCGCTCTATTGTTTAAAG 2880
Db 2821 CAATTGATGCGACCGCTTGAAGCGCTCAACGAATTAACCCGCTCTATTGTTTAAAG 2880
Qy 2881 TGCTGGCGCTGTTCCGTTTTTATCAATCCATCAAAAGAAATTCGATTTCCGGAGGAGAG 2940
Db 2881 TGCTGGCGCTGTTCCGTTTTTATCAATCCATCAAAAGAAATTCGATTTCCGGAGGAGAG 2940
Qy 2941 AGGTCAATCTCCGACATTTGACGCCATTTAGGGCTTTTACGGCGAAACTCCATTTTTCG 3000
Db 2941 AGGTCAATCTCCGACATTTGACGCCATTTAGGGCTTTTACGGCGAAACTCCATTTTTCG 3000
Qy 3001 GAGACTACTTAAACACTGCGGGCAAGAGGAGAGGAGGATCATAAATGCTGAGTGATT 3060
Db 3001 GAGACTACTTAAACACTGCGGGCAAGAGGAGAGGAGGATCATAAATGCTGAGTGATT 3060
```

```
Qy 3061 TAGGCTTTGAAGTTGAATCAGTCAAGAAATGAAGGCTAGCTTAAAGTGCAGAAAGCTGAA 3120
Db 3061 TAGGCTTTGAAGTTGAATCAGTCAAGAAATGAAGGCTAGCTTAAAGTGCAGAAAGCTGAA 3120
Qy 3121 AGAATCAATAAAGCAATCGGTATGATGTCGAATTC 3156
Db 3121 AGAATCAATAAAGCAATCGGTATGATGTCGAATTC 3156

RESULT 2
AAQ81792
ID AAQ81792 standard; DNA; 8478 BP.
XX
XX AAQ81792;
XX AC
XX 25-MAR-2003 (revised)
DT 26-SEP-1995 (first entry)
XX
XX B. subtilis biotin operon and flanking sequences.
XX
XX Biotin operon; recombinant production; dietary additive; animal feed;
XX vitamin supplement; research reagent; ds.
XX
XX Bacillus subtilis.
XX
XX Key Location/Qualifiers
FH terminator 249..291
FT /*tag= k
FT /note= "rho-independent termination site"
FT promoter 324..352
FT /*tag= a
FT /note= "sigma A promoter"
FT misc_feature 355..387
FT /*tag= b
FT /note= "potential regulatory site"
FT CDS 439..1218
FT /*tag= c
FT CDS 1208..2554
FT /*tag= d
FT /product= "bioA"
FT CDS 2544..3713
FT /*tag= e
FT /product= "bioP"
FT CDS 3710..4405
FT /*tag= f
FT /product= "bioD"
FT CDS 4408..5415
FT /*tag= g
FT /product= "bioB"
FT terminator 5423..5462
FT /*tag= h
FT /product= "bioI"
FT CDS 6748..7509
FT /*tag= i
FT /label= ORF 2
FT terminator 7501..7543
FT /*tag= m
FT /note= "rho-independent termination site"
FT CDS 7695
FT /*tag= j
FT /label= ORF 3
XX
XX BP635572-A2.
XX
XX 25-JAN-1995.
XX
XX 13-JUN-1994; 94EP-00108998.
XX 25-JUN-1993; 93US-00084709.
```

PR 06-MAY-1994; 94US-00239430.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX Bower SG, Perkins JB, Pero JG, Yocum RR;
XX WPI; 1995-053684/08.
XX
XX Biotin genes, and constructs derived from Bacillus subtilis - for
PT improved production of recombinant biotin or biotin precursor for use in
PT e.g. dietary supplements.
XX
XX Example II; Fig 14; 75pp; English.
XX
XX AA081792 is the B. subtilis biotin operon and flanking sequences, as part
CC of an expression vector it can be used in the recombinant production of
CC biotin (or biotin precursor protein). The biotin can be used as a dietary
CC additive in animal feeds, and as a vitamin supplement for human
CC consumption. Biotin is also useful as a reagent for research, and
CC diagnostic procedures. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 8478 BP; 2432 A; 1748 C; 2088 G; 2205 T; 0 U; 5 Other;
SQ
Query Match 51.1%; Score 1612.6; DB 2; Length 8478;
Best Local Similarity 73.0%; Pred. No. 0;
Matches 2303; Conservative 0; Mismatches 624; Indels 226; Gaps 9;
QY 2 GATCACAGGTTACGAGCCTTGAAGATTGATCTCGTGTAAACAGAGCGGTTAGACAGAA 61
DB 2524 GATCACAGGTTACGAGCCTTGAAGATTGATCTCGTGTAAACAGAGCGGTTAGACAGAA 2583
QY 62 TGAAGAAGCGCGGTACATCTGTAACCTCGCGTCAATCGATGGAGCGCGGTTCCAGAGA 121
DB 2584 TGAAGAAGCGCGGTACATCTGTAACCTCGCGTCAATCGATGGAGCGCGGTTCCAGAGA 2643
QY 122 GGAATATTGATGGCGAAATCAACGCGTCTCGTCTCTCAACAAATTATTAGGCGTCGCAA 181
DB 2644 GGAATATTGATGGCGAAATCAACGCGTCTCGTCTCTCAACAAATTATTAGGCGTCGCAA 2703
QY 182 CGGATAGAGTTTGTATCGATGAGCCCAACAGAGCAATTCAGCAATTTGGGACAGGAAGCA 241
DB 2704 CGGATAGAGTTTGTATCGATGAGCCCAACAGAGCAATTCAGCAATTTGGGACAGGAAGCA 2763
QY 242 CGCGTTACGTTTAAACGACAGCAATTCGCTGTCGATGAAAGCTAGAAAGAGATTG 301
DB 2764 CGCGTTACGTTTAAACGACAGCAATTCGCTGTCGATGAAAGCTAGAAAGAGATTG 2823
QY 302 CCAGCTTTAAACTGACAGAAGCGCCCTGCTGTTTTCGAGCGGTTACTTTGGCCCAATGTCG 361
DB 2824 CCAGCTTTAAACTGACAGAAGCGCCCTGCTGTTTTCGAGCGGTTACTTTGGCCCAATGTCG 2883
QY 362 GTGTCTTTCACTCTTGCAGAAAAGGAAGATGCTATTTAAGTGAACAGCTCAATCATG 421
DB 2884 GTGTCTTTCACTCTTGCAGAAAAGGAAGATGCTATTTAAGTGAACAGCTCAATCATG 2943
QY 422 CAAGTATCATCGAGCGCTCGGACTTTCTAAGGCTGATACAGTGTGTTTATCGSCATATTG 481
DB 2944 CAAGTATCATCGAGCGCTCGGACTTTCTAAGGCTGATACAGTGTGTTTATCGSCATATTG 3003
QY 482 ATATGAATGATCTTGAAGAACAGCTGGAATGAACACAGCGGTTATCAGCGCGGTTTTATCG 541
DB 3004 ATATGAATGATCTTGAAGAACAGCTGGAATGAACACAGCGGTTATCAGCGCGGTTTTATCG 3063
QY 542 TAAACAGCGGAGTATTACAGATGAGTGAACATCGCCCTCTTGATCAGATCATCTCAC 601
DB 3064 TAAACAGCGGAGTATTACAGATGAGTGAACATCGCCCTCTTGATCAGATCATCTCAC 3123
QY 602 TTGCGAAACGCTATCATCCCTTCGTGCTGTTGATGATGCCACGACAGAGTTTGG 661
DB 3124 TTGCGAAACGCTATCATCCCTTCGTGCTGTTGATGATGCCACGACAGAGTTTGG 3183
QY 662 CGGATTCGGGCAAGAGGACGAGTGAATCTTTGGTGTGTTGTCGCCGACATTTATCGGCA 721
|||||

DB 3184 GCGATTCGGGACAAAGGAAACGAGTGAATACATTGTTGTTTGTCCGACATTTGTTATCGGCA 3243
QY 722 CTTTAAAGCAAGCTGTTTGGCGGAAAGAGGTTTTTGGCGAGGATCAGCGGCTTTCATCG 781
DB 3244 CTTTAAAGCAAGCTGTTTGGCGGAAAGAGGTTTTTGGCGAGGATCAGCGGCTTTCATCG 3303
QY 782 ACTTTTGTGTAACCATGCGCAGAACATTATCTTTCAAAACCGCTATTCCGCCAGCCAGCT 841
DB 3304 ACTTTTGTGTAACCATGCGCAGAACATTATCTTTCAAAACCGCTATTCCGCCAGCCAGCT 3363
QY 842 GTGCGGTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGGAGAAACGACAGCTTT 901
DB 3364 GTGCGGTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGGAGAAACGACAGCTTT 3423
QY 902 TATTTTCTTATATCAGCATGATCAGAACCATGCTGAAGATATGGTATGTTGCTGAAG 961
DB 3424 TATTTTCTTATATCAGCATGATCAGAACCATGCTGAAGATATGGTATGTTGCTGAAG 3483
QY 962 GAGATCACACACCGATTATCTCTGATGCTATGCGGATGCCATATAAACCGTCTTATTG 1021
DB 3484 GAGATCACACACCGATTATCTCTGATGCTATGCGGATGCCATATAAACCGTCTTATTG 3543
QY 1022 CTGAAAACTGCGAGGCAAGGAAATTTATGCTCTGCCATTCGCGCCGCAACCGTTGCGC 1081
DB 3544 CTGAAAACTGCGAGGCAAGGAAATTTATGCTCTGCCATTCGCGCCGCAACCGTTGCGC 3603
QY 1082 CGGTTGAAGCCGGAATTCGAAGCTTTGGGACGAGGTCGAGATCAGGGAATGAGTTATAA 1141
DB 3604 CGGTTGAAGCCGGAATTCGA-----ATTA 3627
QY 1142 AATAAAAAAGCACCCTGAAAAGGTGCTTTTTTGTGATGGTTTTGAACTGTTCTTCTTA 1201
DB 3628 CAATCAGCTGACCAAGTATGGTGTATTTGATCATTTGCTGCAACATTTTCATTCAA 3687
QY 1202 TCTTGATACATATAGAAATAACGTCATTTTATTTTTTATTTTATTTAGTGTGCTGAAAGTGCGT 1261
DB 3688 TCGAAAAGGAGCTGCACATCATTTGAGGGGTTTTTTTGTGACGGGAACGTA----- 3738
QY 1262 TGAAGTGTGGTATGATGTTTTTAAAGTATTGAABACCCCTTAAATTTGTTGCACAGA 1321
DB 3739 -----TACAGAAGTAGGGAACCGTTATATCCACGGGCTTTG- 3776
QY 1322 AAAACCCCATCTGTTAAAGTTATAAGTGACATAAAACAAATAACTAAATAGATGGGGTTTC 1381
DB 3777 -----CTGCCTTATTGAAGACAATAATAGACATGTCGGGTGTA 3816
QY 1382 TTTTAAATATTATGTCCTTAATAGTAGCA TTTTATTCAGATGAAAAATCAAGGTTTTAGT 1441
DB 3817 TAAACCATTTTAAAGCGGGATATCGCGCCATCATCCAGATAGTGTATACAAGTTTGTGTA 3876
QY 1442 GGRCAAGACAAAGTGGAAAAAGTGAGACCAATGTCCTTAGGAAGACGAGTTTATAATAGC 1501
DB 3877 AGATATGTCGAGA----- 3890
QY 1502 TGAATAAGACGGTGTCTCCAAATATTCTTATTAGAAAGCAAAATCTAAAAATTATCTG 1561
DB 3891 -----CCAGTCTTCTCATGAGACATTAACGCTTTTGGCTTCAA 3930
QY 1562 AAAAGGAATAGAAATAGTGAATGGACCAATAATAATGACTAGAGAGAAAGATGAAGA 1621
DB 3931 GGCGCCCTTGCACCATACGTTGACGGGAACTTTGAGGGAAGAGACTGTCAACATGGAAGA 3990
QY 1622 TTGTTTCATGAATTAAGGAACGATATTGGATTAATATGGGATGATGTTAAGGCTATTG 1681
DB 3991 GGTTTTAAAGCCATTGGGGCGGATTAGAGAAAA-----CATGAATGCTTCATCGTAG 4043
QY 1682 GTGTTTATGGCTCTCTTGGTCTGCAGACTGATGGGCCCTATTTCGGATATTGAGATGATG 1741
DB 4044 AAGGTGACGGGGTATTCTGTGCCATTGGGAGAGGACTATTTGTCAGTCATGTCATAA 4103
QY 1742 GTGTCATGTCAACAGAGGAAGCAGAGTTTCAGCCATGAATGGAACACCGGTGAGTGAAGG 1801
DB 4104 -----AAGCGTTGCAGCTTCCCATGATTATTGTTGGCGCTCTCGCTTGGAAACC 4153

Db 1466 AGCTTGGCAGCAGGTGAGATCAGGGAATGAGTTTATAAATATAAATAAAGAACCTGAAA 1525
Qy 1467 |||||||
Qy 1468 |||||||
Db 1469 AGGTGCTCTTTTGTGAGCTTTTGAACCTTTGAACTTCTCTTTCTTATCTTTGATACATATAGAAATA 1221
Db 1470 AGGTGCTCTTTTGTGAGCTTTTGAACCTTTGAACTTCTCTTTCTTATCTTTGATACATATAGAAATA 1585
Qy 1471 |||||||
Qy 1472 ACCTGATCTTTTATTTTATTTTATTTTGTCTGTAAGGTGCTTTGAACTTCTGTAATGATGTTATGT 1281
Db 1473 |||||||
Db 1474 ACCTC-----ATTTTATTTTATTTTGTCTGTAAGGTGCTTTGAACTTCTGTAATGATGTTATGT 1639
Qy 1475 |||||||
Qy 1476 GTTTTAAAGTATTGAAAAACCTTAAAAATTTGGTTCACAGAAAAACCCCATCTGTTTAAAGT 1341
Db 1477 |||||||
Db 1478 GTTTTAAAGTATTGAAAAACCTTAAAAATTTGGTTCACAGAAAAACCCCATCTGTTTAAAGT 1699
Qy 1479 |||||||
Db 1480 TATAAGTGACTAAACAAATAAATACTAAATAGATGGGGTTTCTTTTAAATAATTAATGTCCTA 1401
Db 1481 TATAAGTGACTAAACAAATAAATACTAAATAGATGGGGTTTCTTTTAAATAATTAATGTCCTA 1759
Qy 1482 ATAGTAGCATTTATTCAGATGAAAAATCAAGGTTTATAGTCGACAGCAAAAAAGTGAA 1461
Db 1483 |||||||
Db 1484 ATAGTAGCATTTATTCAGATGAAAAATCAAGGTTTATAGTCGACAGCAAAAAAGTGAA 1819
Qy 1485 |||||||
Qy 1486 AAGTGAGACCATGTGCTTAGGAAGCAGGTTTATTAATAGCTGAATAAGAACGGTCTCTC 1521
Db 1487 |||||||
Db 1488 AAGTGAGACCATGTGCTTAGGAAGCAGGTTTATTAATAGCTGAATAAGAACGGTCTCTC 1879
Qy 1489 |||||||
Db 1490 CAAATATTTCTTTATTTAGAAAAAGCAAACTCTAAATTTATCTGAAAAAGGGAATGAGAATAGTG 1581
Db 1491 CAAATATTTCTTTATTTAGAAAAAGCAAACTCTAAATTTATCTGAAAAAGGGAATGAGAATAGTG 1939
Qy 1492 AATGACCAATAATAATGACCTAGAGNAGAAAGATGAAGTTTTCATGAAATTAAGGA 1641
Db 1493 |||||||
Db 1494 AATGACCAATAATAATGACCTAGAGNAGAAAGATGAAGTTTTCATGAAATTAAGGA 1999
Qy 1495 |||||||
Qy 1496 CGAATATTTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGT 1701
Db 1497 |||||||
Db 1498 CGAATATTTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGT 2059
Qy 1499 |||||||
Qy 1500 CGTCAGACTGATGGGCCCTATTTCGGATATTCAGATGATGTTGTCATGTCACAGAGGAA 1761
Db 1501 CGTCAGACTGATGGGCCCTATTTCGGATATTCAGATGATGTTGTCATGTCACAGAGGAA 2119
Qy 1502 GCAGAGTTCCAGCCATGAATGACACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGC 1821
Db 1503 |||||||
Db 1504 GCAGAGTTCCAGCCATGAATGACACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGC 2179
Qy 1505 |||||||
Qy 1506 GAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGT 1881
Db 1507 |||||||
Db 1508 GAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGT 2239
Qy 1509 |||||||
Qy 1510 CAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAA 1941
Db 1511 |||||||
Db 1512 CAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAA 2299
Qy 1513 |||||||
Qy 1514 ACTGCTAAATCGGTAGAGCCCAACGCTTCCACGATCGATTTCTGCCCTTATCGTAGAA 2001
Db 1515 |||||||
Db 1516 ACTGCTAAATCGGTAGAGCCCAACGCTTCCACGATCGATTTCTGCCCTTATCGTAGAA 2359
Qy 1517 |||||||
Qy 1518 GAGCTGTTTGAATATGCAAGGCAATGGCGTAAATATTCTGTGTGCAAGGACCCGACAACTTT 2061
Db 1519 |||||||
Db 1520 GAGCTGTTTGAATATGCAAGGCAATGGCGTAAATATTCTGTGTGCAAGGACCCGACAACTTT 2419
Qy 1521 |||||||
Qy 1522 CTACATCTTTGACTGTACAGGTAGCAATGAGGTGCCATGTTGATGTTGCTGCATCAT 2121
Db 1523 |||||||
Db 1524 CTACATCTTTGACTGTACAGGTAGCAATGAGGTGCCATGTTGATGTTGCTGCATCAT 2479
Qy 1525 |||||||
Qy 1526 CGCATCTGTTATACGACGAGCGCTTCGCTCTTAATCTGAAGCAGTTAAGCAATCAGATCTT 2181
Db 1527 |||||||
Db 1528 CGCATCTGTTATACGACGAGCGCTTCGCTCTTAATCTGAAGCAGTTAAGCAATCAGATCTT 2539
Qy 1529 |||||||
Qy 1530 CCTTCAGGTTATGACCATCTGTGCGAGTTTCGTAATGTCGTGTCACCTTTCCGACTCTGAG 2241
Db 1531 |||||||

Db 2540 CCTTCAGGTTATGACCATCTGTGCCAGTTCGTAAATGTCGTCAACTTTCGACTCTGAG 2599
Qy 2541 |||||||
Qy 2542 AAACCTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAAGAGTGGACAGAACGACAC 2301
Db 2543 |||||||
Db 2544 AAACCTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAAGAGTGGACAGAACGACAC 2659
Qy 2545 |||||||
Qy 2546 GGATATATAGTGGATGTCGTAACAAACGCATACCATTTTGAAT 2342
Db 2547 |||||||
Db 2548 GGATATATAGTGGATGTCGTAACAAACGCATACCATTTTGAAT 2700
RESULT 4
ADT25766
ID ADT25766 standard; DNA; 7350 BP.
XX
AC ADT25766;
XX
DT 13-JAN-2005 (first entry)
XX
DE Vitamin B12 preparation 7350bp vector.
XX
KW vitamin B12; transgenic; vector; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT promoter 9..18
FT /*tag= a
FT /label= xyl-A promoter
FT CDS 27..953
FT /*tag= b
FT /gene= "cbix B. megaterium"
FT rep_origin 1123..1173
FT /*tag= c
FT /label= oriE.coli (pBR322)
FT CDS 1906..2766
FT /direction= Counterclockwise
FT /*tag= d
FT /gene= "Amp-resistance gene"
FT rep_origin 3611..3686
FT /*tag= e
FT /label= oriB.megabacterium (pB194ts)
FT CDS 4008..4615
FT /direction= Counterclockwise
FT /*tag= f
FT /label= repF
FT CDS 6399..7169
FT /direction= Counterclockwise
FT /*tag= g
FT /gene= "Neo-resistance gene"
XX
W02004090138-A1.
XX
PD 21-OCT-2004.
XX
PF 10-MAR-2004; 2004WO-EP002430.
XX
PR 12-APR-2003; 2003DB-01016888.
XX
PA (BADI) BASF AG.
XX
PI Barg H, Jahn D;
XX
XX WPI; 2004-757985/74.
DR Preparation of Vitamin B12 by culturing microorganisms that have been
PT transformed with a vector containing a cobalt chelataase gene and
PT inducible promoter.
XX
PS Claim 1; SEQ ID NO 1; 57pp; German.
XX
CC The invention relates to a novel method for the preparation of vitamin
CC B12. The method involves culturing a microorganism that has been

CC transformed with a vector that includes a 7350bp sequence ADT25766. The
CC invention further relates to a vector that contains the 7350bp sequence;
CC and genetically modified microorganisms that contain said vector. This
CC polynucleotide sequence represents the 7350bp vector used in the
CC preparation of vitamin B12 of the invention.

XX SQ Sequence 7350 BP; 2397 A; 1340 C; 1580 G; 2033 T; 0 U; 0 Other;

Query Match 38.0%; Score 1200.4; DB 13; Length 7350;
Best Local Similarity 99.0%; Pred. No. 1.5e-288;
Matches 1233; Conservative 0; Mismatches 1; Indels 12; Gaps 2;

Qy 1102 AGCTGGGCGAGCGTCCGAGATCAGGGAATCAGTTTATAAAATAAAAAAGCACCTGAAA 1161
Db 5925 AGCTTGGGCTGCAGGTCGAGATCAGGGAATCAGTTTATAAAATAAAAAAGCACCTGAAA 5984
Qy 1162 AGGTGCTCTTTTGTGATGGTTTGAACCTGTTCTTCTATCTTATCTTGATACATATAGAATA 1221
Db 5985 AGGTGCTCTTTTGTGATGGTTTGAACCTGTTCTTCTATCTTATCTTGATACATATAGAATA 6044
Qy 1222 ACGTCATTTTATTTTATTTTATTTAGTTCGTGAAAGTGGTGTGAAGTCTTGGTATGTATGT 1281
Db 6045 ACGTC-----ATTTTATTTTATTTTATTTAGTTCGTGAAAGTGGTGTGAAGTCTTGGTATGTATGT 6098
Qy 1282 GTTTTAAAGTATTGAAACCCCTTAAATTTGGTTCACAGAAAAACCCCATCTGTTTAAAGT 1341
Db 6099 GTTTTAAAGTATTGAAACCCCTTAAATTTGGTTCACAGAAAAACCCCATCTGTTTAAAGT 6158
Qy 1342 TATAAGTGACTAAACAAATAACTAAATAGATGGGGGTTTCTTTTAATATATTATGTGCTTA 1401
Db 6159 TATAAGTGACTAAACAAATAACTAAATAGATGGGGGTTTCTTTTAATATATTATGTGCTTA 6218
Qy 1402 ATAGTAGCATTTATTCAGATGAAAAATCAAGGTTTTAGTGGCAACAGAAAAAGTGGAA 1461
Db 6219 ATAGTAGCATTTATTCAGATGAAAAATCAAGGTTTTAGTGGCAACAGAAAAAGTGGAA 6278
Qy 1462 AAGTGAGACCATG-----TGCTTAGGAAGACGAGTTATTAAATAGCTGAATAAGAACGGT 1515
Db 6279 AAGTGAGACCATGAGCTTATGCTTAGGAAGACGAGTTATTAAATAGCTGAATAAGAACGGT 6338
Qy 1516 GCTCTCCAAATATCTTTATTAGAAAAAGCAAAATCTAAAAATTTATCTGAAAAAGGGAATGAGA 1575
Db 6339 GCTCTCCAAATATCTTTATTAGAAAAAGCAAAATCTAAAAATTTATCTGAAAAAGGGAATGAGA 6398
Qy 1576 ATAGTGAATGACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAATTT 1635
Db 6399 ATAGTGAATGACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAATTT 6458
Qy 1636 AAGGAACGAATATTGGGATAAATATGGGATGATGTTAAGGCTATTGGTGTTTATGGCTCT 1695
Db 6459 AAGGAACGAATATTGGGATAAATATGGGATGATGTTAAGGCTATTGGTGTTTATGGCTCT 6518
Qy 1696 CTTGGTCGTGAGACTGATGGGCCCTATTTCGGAATATGAGATGATGTGTGTCATGTCACAA 1755
Db 6519 CTTGGTCGTGAGACTGATGGGCCCTATTTCGGAATATGAGATGATGTGTGTCATGTCACAA 6578
Qy 1756 GAGGAACGAGGTTTACCGATGAATGACACACCGGTGAGTGGAGGGAAGTGAATTTT 1815
Db 6579 GAGGAACGAGGTTTACCGATGAATGACACACCGGTGAGTGGAGGGAAGTGAATTTT 6638
Qy 1816 GATAGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACA 1875
Db 6639 GATAGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACA 6698
Qy 1876 CATGGTCAATTTTCTCTATTGTCGGAATTATGATTCAGGTGATATCTTAGAGAAAGTG 1935
Db 6699 CATGGTCAATTTTCTCTATTGTCGGAATTATGATTCAGGTGATATCTTAGAGAAAGTG 6758
Qy 1936 TATCAAACTGTAATCGGTAGAGCCCAACGTTCCACGATCGATTGTCGCTTATC 1995
Db 6759 TATCAAACTGTAATCGGTAGAGCCCAACGTTCCACGATCGATTGTCGCTTATC 6818
Qy 1996 GTAGAAGAGCTGTTTGAATATGCAAGGCAAAATGGCGTAATATTTCGTGTGCAAGGACCGACA 2055

Db 6819 GTAGAAGAGCTGTTTGAATATGATGAGGCAAAATGCGTAATATTTCGTGTGCAAGGACCGACA 6878
Qy 2056 ACATTTCTACCATCCTTGAAGTGTACAGGTAGCAATGCGAGTGCCATGTTGATTGGTCTG 2115
Db 6879 ACATTTCTACCATCCTTGAAGTGTACAGGTAGCAATGCGAGTGCCATGTTGATTGGTCTG 6938
Qy 2116 CATCATCGCATCTGTTATATACGACGAGCGCTTCTTAAGCAGTTAAAGCAATCA 2175
Db 6939 CATCATCGCATCTGTTATATACGACGAGCGCTTCTTAAGCAGTTAAAGCAATCA 6998
Qy 2176 GATCTTCTTTCAGGTTATGACCATCTGTGTCAGTTTGTAAATGTCGTCAACTTCGAC 2235
Db 6999 GATCTTCTTTCAGGTTATGACCATCTGTGTCAGTTTGTAAATGTCGTCAACTTCGAC 7058
Qy 2236 TCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAA 2295
Db 7059 TCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAA 7118
Qy 2296 CGACACGATATATAGTGGATGTGTCAAAACGCATACCATTTTGAA 2341
Db 7119 CGACACGATATATAGTGGATGTGTCAAAACGCATACCATTTTGAA 7164

RESULT 5
AEB16616
ID AEB16616 standard; DNA; 8152 BP.
XX AEB16616;
AC AEB16616;
XX 25-AUG-2005 (first entry)
XX Plasmid pAN213ban.
DE
XX
KW Oxidative stress; DNA damage; protein secretion; amyO; ds; plasmid;
KW amyase.
OS Bacillus subtilis.
OS Synthetic.
OS Unidentified.
XX
PN WO2005056799-A1.
XX
PD 23-JUN-2005.
XX
PF 10-DEC-2004; 2004WO-DK000859.
XX
PR 10-DEC-2003; 2003DK-00001824.
XX
PA (NOVO) NOVOZYMES AS.
PI Nielsen AK, Rasmussen MD;
XX
XX WPI; 2005-445177/45.
XX
PT Novel progeny cell derived from parent cell, and comprising gene encoding
PT metallo regulated gene A MrgA protein or its functional homolog and/or
PT DNA segment operably linked with encoding gene, useful for producing
PT protein of interest.
XX
PS Example 4; SEQ ID NO 16; 49pp; English.
XX
CC The invention relates to a progeny cell derived from a parent cell, and
CC comprising a gene encoding metallo regulated gene A (MrgA) protein or its
CC functional homolog and/or DNA segment operably linked with the encoding
CC gene, where the gene and/or DNA segment is manipulated with respect to
CC the parent cell, or being mutated with respect to the parent cell, where
CC the cell produces greater amounts of MrgA protein or its functional
CC homolog than the parent cell. Also included are enhancing secretion of a
CC protein of interest (involving expressing the protein in the cell),
CC preparation of the progeny cell and use of the MrgA-protein or its
CC functional homologue for enhancing production of a protein by
CC manipulating or mutating a cell to express greater amounts of MrgA-

comprising a gene encoding metallo regulated gene A (MrGA) protein or its functional homolog and/or DNA segment operably linked with the encoding gene, where the gene and/or DNA segment is manipulated with respect to the parent cell, or being mutated with respect to the parent cell, where the cell produces greater amounts of MrGA protein or its functional homolog than the parent cell. Also included are enhancing secretion of a protein of interest (involving expressing the protein in the cell), preparation of the progeny cell and use of the MrGA-protein or its functional homolog for enhancing production of a protein by manipulating or mutating a cell to express greater amounts of MrGA-protein or its functional homologue than the non-manipulated or non-mutated cell. MrGA, a protein involved in oxidative stress response, has been demonstrated to increase the secretion of proteins when over expressed. The progeny cell is useful for producing a protein of interest, where the protein of interest is a protease, lipase, cutinase, an amylase, galactosidase, pullulanase, cellulase, glucose isomerase, protein disulfide isomerase, cyclodextrin gluconotransferase (CGR'ase), phytase, glucose oxidase, glucosyl transferase, lactase, bilirubin oxidase, xylanase, antigenic microbial or protozoan protein, bacterial protein toxin, microbial surface protein, or a viral protein. MrGA or its functional homolog is useful for enhancing production of a protein. The present sequence is the plasmid pDG268neo which is unable to replicate in bacillus subtilis (source of the MrGA gene) but which contains a chloramphenicol resistance marker, cam, (and restriction sites for cloning) flanked by the 5' and 3' regions of the B. subtilis amyE gene. The plasmid is used for introducing an MrGA/cam marker expression cassette into the amyE locus.

Query Match 37.8%; Score 1192.8; DB 14; Length 8644;
Best Local Similarity 98.9%; Pred. No. 1.3e-286;
Matches 1224; Conservative 0; Mismatches 7; Indels 7; Gaps 2;

Qy	1105	TTGGGACGAGTCGACATCAGGGAATGAGTTTATAAAATAAAAGACCTGAAAGG	1164
Db	6664	TGGAGGAGTATGGGGAGATCAGGGAATGAGTTTATAAAATAAAAGACCTGAAAGG	6605
Qy	1165	TGCTCTTTTGTAGTGTGTTGAACTGCTCTTCTTATCTTGATACATATAGAAATACG	1224
Db	6604	TGCTCTTTTGTAGTGTGTTGAACTGCTCTTCTTATCTTGATACATATAGAAATACG	6545
Qy	1225	TCATTTTATTTTATTTAGTTGCTGAAAGGTCGTTGAAAGTGTGCGTATGATGTGTT	1284
Db	6544	TC-----ATTTTATTTTATTTAGTTGCTGAAAGGTCGTTGAAAGTGTGCGTATGATGTGTT	6491
Qy	1285	TTAAAGTATTGAAACCCCTTAAATTTGCTGACAGAAAAACCCCATCTGTTAAAGTTAT	1344
Db	6490	TTAAAGTATTGAAACCCCTTAAATTTGCTGACAGAAAAACCCCATCTGTTAAAGTTAT	6431
Qy	1345	AAGTGACTAAACAATAAATAAGATAGATGGGGTTCTTTTAAATATTATGTCCTTAATA	1404
Db	6430	AAGTGACTAAACAATAAATAAGATAGATGGGGTTCTTTTAAATATTATGTCCTTAATA	6371
Qy	1405	GTAGCATTTATTCAGATGAAAAATCAAGGGTTTTAGTGGACACAAAAAGTGGAAAAAG	1464
Db	6370	GTAGCATTTATTCAGATGAAAAATCAAGGGTTTTAGTGGACACAAAAAGTGGAAAAAG	6311
Qy	1465	TGAGACCATG-TGCTTAGGAAGACGAGTTTAAATAGCTGAAATGAAGAACGGTCTCTCCA	1523
Db	6310	TGAGACCATGATGCTTAGGAAGACGAGTTTAAATAGCTGAAATGAAGAACGGTCTCTCCA	6251
Qy	1524	AATATTCCTTATTAGAAAAACAATCTAAATTTATCTGAAAAAGGAATGAGATAGTAA	1583
Db	6250	AATATTCCTTATTAGAAAAACAATCTAAATTTATCTGAAAAAGGAATGAGATAGTAA	6191
Qy	1584	TGGACCAATAATATGACTAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAGGAACG	1643
Db	6190	TGGACCAATAATATGACTAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAGGAACG	6131
Qy	1644	AATATTCGATAATATCGGGAGTATGTTAAGGCTATTGGTGTGTTATGGCTCTCTTGTGCG	1703
Db	6130	AATATTCGATAATATCGGGAGTATGTTAAGGCTATTGGTGTGTTATGGCTCTCTTGTGCG	6071

Qy	1704	TCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGTATGTCAACACAGAGGAAGC	1763
Db	6070	TCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGTATGTCAACACAGAGGAAGC	6011
Qy	1764	AGAGTTTCAGCCATGAATGGACAAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCGA	1823
Db	6010	AGAGTTTCAGCCATGAATGGACAAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCGA	5951
Qy	1824	AGAGATTCTACTAGATTATGCACTCTCAGGTGGGAATCAGATTGGCCGCTTACACATGTGCA	1883
Db	5950	AGAGATTCTACTAGATTATGCACTCTCAGGTGGGAATCAGATTGGCCGCTTACACATGTGCA	5891
Qy	1884	ATTTTCTCTTATTTTGGCCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTATCAAAAC	1943
Db	5890	ATTTTCTCTTATTTTGGCCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTATCAAAAC	5831
Qy	1944	TGCTAAATCGGTAGAGACCCCAAAAGTTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGA	2003
Db	5830	TGCTAAATCGGTAGAGACCCCAAAAGTTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGA	5771
Qy	2004	GCTGTTTGAATATGACGGCAAAATGGCGTAAATATTCGTGTGCAAGGACCGACAACATTTCT	2063
Db	5770	GCTGTTTGAATATGACGGCAAAATGGCGTAAATATTCGTGTGCAAGGACCGACAACATTTCT	5711
Qy	2064	ACCATCTTGTGACTGTACAGGTAGCAATGGCAGGTGCCATCTTGTATGCTGTCATCATCG	2123
Db	5710	ACCATCTTGTGACTGTACAGGTAGCAATGGCAGGTGCCATCTTGTATGCTGTCATCATCG	5651
Qy	2124	CATCTGTTTATACGACGAGCGCTTCCGCTTTAACTGAAGCAGTGTAAAGCAATCAGATCTTCC	2183
Db	5650	CATCTGTTTATACGACGAGCGCTTCCGCTTTAACTGAAGCAGTGTAAAGCAATCAGATCTTCC	5591
Qy	2184	TTCAGGTTATGACCATCTGTGCCAGTTTCGTAATGTCTGCTCAACTTCCGACTCTGAGAA	2243
Db	5590	TTCAGGTTATGACCATCTGTGCCAGTTTCGTAATGTCTGCTCAACTTCCGACTCTGAGAA	5531
Qy	2244	ACTTCTGGAAATCCCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGG	2303
Db	5530	ACTTCTGGAAATCCCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGG	5471
Qy	2304	ATATATAGTGGATGTGTCAAAAACGCAATACCATTTTGA 2341	
Db	5470	ATATATAGTGGATGTGTCAAAAACGCAATACCATTTTGA 5433	

RESULT 7
ABA06372/c
ID ABA06372 standard; DNA; 13129 BP.
XX
AC ABA06372;
XX
XX 14-JUL-2005 (first entry)
XX
DE Novel bacterial host cell-related plasmid pCLO1154 DNA SeqID8.
XX gene expression; protein production; chromosome; genome; cell culture;
XX antibiotic-resistance; ds.
XX Unidentified.
XX
XX WO2005042750-A1.
XX
PD 12-MAY-2005.
XX
XX 29-OCT-2004; 2004WO-DK000750.
XX
XX 31-OCT-2003; 2003DK-00001624.
XX
XX (NOVO) NOVOZYMES AS.
XX
XX Olsen C, Rasmussen MD;
XX

DR WPI; 2005-346878/35.
XX New bacterial host cell comprising at least two copies of an
PT amplification unit in its genome, useful for increased production of a
PT protein with several copies of a gene of interest stably chromosomally
XX integrated.
XX Example; SEQ ID NO 8; 53pp; English.
XX This invention relates to a novel bacterial host cell comprising at least
CC two copies of an amplification unit in its genome, where two or more
CC amplified chromosomal copies of the gene of interest are produced. The
CC amplification unit comprises at least one copy of a gene of interest and
CC an expressible conditionally essential gene, where the conditionally
CC essential gene is either promoterless or transcribed from a heterologous
CC promoter having an activity substantially lower than the endogenous
CC promoter of the conditionally essential gene, and where the conditionally
CC essential gene if not functional would render the cell auxotrophic for at
CC least one specific substance or unable to utilize one or more specific
CC sole carbon source. The bacterial host cell is useful for producing of a
CC protein having several copies of a gene of interest stably chromosomally
CC integrated, without leaving antibiotic resistance marker genes in the
CC strains. The present sequence is that of a plasmid DNA sequence which was
CC used in the exemplification of the invention.
SQ Sequence 13129 BP; 3491 A; 3037 C; 3009 G; 3591 T; 0 U; 1 Other;

Query Match 37.8%; Score 1192.8; DB 14; Length 13129;
Best Local Similarity 98.9%; Pred. No. 1.5e-286;
Matches 1224; Conservative 0; Mismatches 7; Indels 7; Gaps 2;

QY 1105 TTGGGCACAGCTCGAGATCAGGGAATGAGTTTATAAAATAAAAAAGCACCTGAAAGG 1164
DB 11149 TGGAGGAGTATGGGGAGATCAGGGAATGAGTTTATAAAATAAAAAAGCACCTGAAAGG 11090
QY 1165 TGTCTTTTTCATGGTTTGAACACTGTTCTTTCTTATCTTGTATACATATAGAAATACG 1224
DB 11089 TGTCTTTTTCATGGTTTGAACACTGTTCTTTCTTATCTTGTATACATATAGAAATACG 11030
QY 1225 TCATTTTATTTTATTTTATGTTCTGTAAGGTGGGTTGAGTGTGTTATGTTGTT 1284
DB 11029 TC-----ATTTTATTTTATGTTCTGTAAGGTGGGTTGAGTGTGTTATGTTGTT 10376
QY 1285 TTAAGTATTGAAAACCCCTTAAATTTGTTGACAGAAAACCCCATCTGTTAAAGTTAT 1344
DB 10975 TTAAGTATTGAAAACCCCTTAAATTTGTTGACAGAAAACCCCATCTGTTAAAGTTAT 10916
QY 1345 AAGTGACTAAACAAATACTAAATAGATGGGGTTTCTTTTAAATATTATGTTCTTAATA 1404
DB 10915 AAGTGACTAAACAAATACTAAATAGATGGGGTTTCTTTTAAATATTATGTTCTTAATA 10856
QY 1405 GTAGCATTTATTCAGATGAAAATCAAGGTTTATGTTGACAGAAAATAGTGGAAAAG 1464
DB 10855 GTAGCATTTATTCAGATGAAAATCAAGGTTTATGTTGACAGAAAATAGTGGAAAAG 10796
QY 1465 TGAGACCATG-TGCTTAGGAAGACAGATTATTAATAGCTGTAATGAACGGTCTCTCCA 1523
DB 10795 TGAGACCATGATCTTAGGAAGACAGATTATTAATAGCTGTAATGAACGGTCTCTCCA 10736
QY 1524 AATATTTCTTATTAGAAAAGCAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGA 1583
DB 10735 AATATTTCTTATTAGAAAAGCAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGA 10676
QY 1584 TGGACCAATAATATGACTAGAGAAGAAAGTGAAGTTGTTTCATGAAATTAAGGAACG 1643
DB 10675 TGGACCAATAATATGACTAGAGAAGAAAGTGAAGTTGTTTCATGAAATTAAGGAACG 10616
QY 1644 AATATTGATAAATATGCGGATGATGTTAAGCTATTGCTGTTTATGGCTCTCTTTGGTCG 1703
DB 10615 AATATTGATAAATATGCGGATGATGTTAAGCTATTGCTGTTTATGGCTCTCTTTGGTCG 10556
QY 1704 TCAGACTGATGGGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACACAGAGGAAGC 1763

DB 10555 TCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCAACACAGAGGAAGC 10496
QY 1764 AGAGTTTCAGGCATGAATGGACAAACCGGTGAGTGAAGGTGAAGTGAATTTTGATAGCGA 1823
DB 10495 AGAGTTTCAGGCATGAATGGACAAACCGGTGAGTGAAGGTGAAGTGAATTTTGATAGCGA 10436
QY 1824 AGAGATTCTACTAGATTATGCACTCTCAGGTGGGAATCAGATTGGCCGCTTACACATGGTCA 1883
DB 10435 AGAGATTCTACTAGATTATGCACTCTCAGGTGGGAATCAGATTGGCCGCTTACACATGGTCA 10376
QY 1884 ATTTTCTCTATTTTGGCGAATTATGATTCAGGTGGATATCTAGAGAAAGTGTATCAAAC 1943
DB 10375 ATTTTCTCTATTTTGGCGAATTATGATTCAGGTGGATATCTAGAGAAAGTGTATCAAAC 10316
QY 1944 TGCTAAATCCGTAGAACCCCAAGGTTCCACAGATGCCATTGTGCCCTTATCGTAGAAGA 2003
DB 10315 TGCTAAATCCGTAGAACCCCAAGGTTCCACAGATGCCATTGTGCCCTTATCGTAGAAGA 10256
QY 2004 GCTGTTTGAATATGCAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACAACATTTCT 2063
DB 10255 GCTGTTTGAATATGCAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACAACATTTCT 10196
QY 2064 ACCATCTTGAATATGCAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACAACATTTCT 2123
DB 10195 ACCATCTTGAATATGCAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACAACATTTCT 10136
QY 2124 CATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCC 2183
DB 10135 CATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCC 10076
QY 2184 TTCAGGTTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCC 2243
DB 10075 TTCAGGTTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCC 10016
QY 2244 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGAATTCAGAGTGGAGACGACACCG 2303
DB 10015 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGAATTCAGAGTGGAGACGACACCG 9956
QY 2304 ATATATAGTGGATGTGCAAAACGACATACCATTTTGAA 2341
DB 9955 ATATATAGTGGATGTGCAAAACGACATACCATTTTGAA 9918

RESULT 8
ADT48504
ID ADT48504 standard; cDNA; 1170 BP.
XX
AC ADT48504;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #23255.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.
XX
XX Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
PA

(SLAT//) SLATER S C.
(CHEN//) CHEN X.
(GOLD//) GOLDMAN B S.

Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

Claim 1: SEQ ID NO 46942: 122pp; English.

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polynucleotide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 1170 BP; 321 A; 256 C; 293 G; 300 T; 0 U; 0 Other;

34.3%; Score 1081; DB 13; Length 1170;

Local Similarity Match

Local Similarity 100.0%, Id. NO. 4.3E-233,
ches 1081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

22 TTTGAAGATTGATTCCTGCTTAAACGAGCGGTTAGACAGAATGAAGAAGCCGCGGTACAT 81
|||||
1 TTTGAAGATTGATTCCTGCTTAAACGAGCGGTTAGACAGAATGAAGAAGCCGCGGTACAT 60

82 CGTAA CCTGCCGGTCAATGGATGGAGCGCCGGTTCCAGAGAGGAAATATGTGATGGCGGAAAT 141

61 CGTAA CCTGCGGTCAA TGGATGGAGCGCCGGTTCAGAGAGGAA TATTGATGGCGAAAT 120

142 CAAACGCTCTGGTCCTCAAA CAATTATTAGGGCTCGCAAGCGATAGACGTTTGATCGAT 201

121 CAAACGGTCTGGTCTCTCAAA CAATTATTTAGGGCTCGCAAGCGATAGACGTTTGATCGAT 180

202 GCAGCCCAACAGCATTGCAGCAATTTGGGACAGGAAGCAGCGGTTCA CGTTTACGACA 261

181 GCAGCCCAACAGCATTTGCAGCAATTTTGGGACAGGAAGCAGCGGTTACGTTTACGACA 240

262 GGCAATTCCGGTCTGGCATGAAAAGCTAGAAAAGAAAGATTGCCAGCTTTAAACTGACAGAA 321

241 GGCAATTTCGGTCTGGCATGAAAAGCTAGAAAAGAAGATTGCCAGCTTTAAACTGACAGAA 300

322 GCGGCCCTGCTGTTTTCGAGCGGTTACTTGGCCAAATGTCGGTGTCTTTTCATCCTTGCCA 381

301 GCGGCCCTGCTGTTTTCGAGCGGTA CTTGGCCAATGTCGGTGTCTTTCA TCCTTGCCA 360

382 GAAAAGGAAGATGTCTATTTTAAGTGACCACTCAATCATGCAAGTATGATCGACGGCTGC 441

361 GAAAAGGAAGATGTCA TTTTAA GTGAC CAGCTCAA TCATGCA AGTATGATCGACGCGCTGC 420

Qy	442	CGACTTCTTAAGCGCTGATACAGTGTGTTTATCGCGCATTTATGATATGAATCATCTTGAAAAAC	501
Db	421	CGACTTCTTAAGCGCTGATACAGTGTGTTTATCGCGCATTTATGATATGAATCATCTTGAAAAAC	480
Qy	502	AAGCTGAATGAAACACAGACGCTTATCAGCGCGGTTTTATCGTAAACAGACCGGAGTATTTCAGC	561
Db	481	AAGCTGANTGAAACACAGCGTTATCAGCGCGGTTTTATCGTAAACAGACCGGAGTATTTCAGC	540
Qy	562	ATGGATGGCAAAATCGCCCTCTTGATCAGATCATCTCACTTCGCGAAACGCTATCATGCC	621
Db	541	ATGGATGGCAAAATCGCCCTCTTGATCAGATCATCTCACTTCGCGAAACGCTATCATGCC	600
Qy	622	TTGCTGGTGGTTCATGATGATGCGCCACGCAACGAGGATTTGGCGGATTCGGGACACGAAACG	681
Db	601	TTGCTGGTGGTTCATGATGATGCGCCACGCAACGAGGATTTGGCGGATTCGGGACACGAAACG	660
Qy	682	AGTGAATACATTTGGTGTGTTCTCCCGACATTTTATCGGCACCTTTAAGCAAAAGCTGTTGGC	741
Db	661	AGTGAATACATTTGGTGTGTTCTCCCGACATTTTATCGGCACCTTTAAGCAAAAGCTGTTGGC	720
Qy	742	CGCGAAGGAGGTTTTGCGGCAGGATCAGCGGCTCTTCATCGACTTTTTCGTGAAACCATGCC	801
Db	721	CGCGAAGGAGGTTTTGCGGCAGGATCAGCGGCTCTTCATCGACTTTTTCGTGAAACCATGCC	780
Qy	802	AGAACATTTATCTTTCAAACCGCTATTTCGCGACGACGCTGCGGCTGCTCAGAGGCT	861
Db	781	AGAACATTTATCTTTCAAACCGCTATTTCGCGACGACGCTGCGGCTGCTCAGAGGCT	840
Qy	862	TTCAACATCATTTCAAGCCAGCAGCGGAAACGACAGCTTTTATTTCTTATATCAGCATG	921
Db	841	TTCAACATCATTTCAAGCCAGCAGCGGAAACGACAGCTTTTATTTCTTATATCAGCATG	900
Qy	922	ATCAGAACCAAGTCTGAAGAATATGGGTATTATGGTGAAGGAGATCACACCGATTATT	981
Db	901	ATCAGAACCAAGTCTGAAGAATATGGGTATTATGGTGAAGGAGATCACACCGATTATT	960
Qy	982	CTGTGATCATTTGGCGATGCCCATATAAACCGTCTCTATTTGCTGAAAAACTGCAAGGCGAAG	1041
Db	961	CTGTGATCATTTGGCGATGCCCATATAAACCGTCTCTATTTGCTGAAAAACTGCAAGGCGAAG	1020
Qy	1042	GGAAATTTATGCTCTGTCATTTGCGCGCCCAACCGTTGCGCGGTTGAAAGCCCGATTTCGA	1101
Db	1021	GGAAATTTATGCTCTGTCATTTGCGCGCCCAACCGTTGCGCGGTTGAAAGCCCGATTTCGA	1080
Qy	1102	A 1102	
Db	1081	A 1081	
RESULT 9			
AEA43211			
ID	AEA43211 standard; DNA; 4334 BP.		
XX	AEA43211;		
XX	25-AUG-2005 (first entry)		
DT	M. smegmatis perhydrolase expression vector pAH505.		
XX	perhydrolase; bleaching; paper; pulp; cosmetics; vector; db.		
XX	Mycobacterium smegmatis.		
OS	Synthetic.		
XX	WO2005056782-A2.		
FN	23-JUN-2005.		
XX	03-DEC-2004; 2004WO-US040438.		
PF	03-DEC-2003; 2003US-0526764P.		
XX	(GEMV) GENENCOR INT INC.		

PA	(PROC) PROCTER & GAMBLE CO.	
XX	Amin NS, Boston MG, Bott RR, Cervin MA, Concar EM, Gustwiller ME;	
PI	Jones BE, Liebecon K, Miracle GS, Oh H, Poulouse AU, Ramer SW;	
PI	Scheibel JJ, Weyler W, Whited GM;	
XX		
DR	WPI; 2005-458588/46.	
XX		
XX	New isolated perhydrolase having preset perhydrolysis to hydrolysis	
PT	ratio, useful in cleaning, bleaching and disinfecting applications e.g.	
PT	paper pulp bleaching.	
XX		
PS	Example 6; SEQ ID NO 137; 523pp; English.	
XX		
CC	The invention relates to an isolated perhydrolase (I), exhibiting a	
CC	perhydrolysis to hydrolysis ratio that is greater than 1. A composition	
CC	containing (I) is useful for cleaning. (I) is useful in cleaning,	
CC	bleaching and disinfecting applications such as paper and pulp bleaching,	
CC	fabric and garment cleaning, hard surface cleaning and personal care	
CC	applications (e.g. oral care, hair care, and skin care). (I) is useful in	
CC	applications where peracid hydrolysis is useful, such as in the clean up	
CC	of peracids. (I) enables effective cleaning, bleaching and disinfecting	
CC	over broad pH and temperature ranges. The present sequence represents a	
CC	M. smegmatis perhydrolase expression vector.	
XX		
SQ	Sequence 4334 BP; 1354 A; 760 C; 1028 G; 1192 T; 0 U; 0 Other;	
Query Match 27.5%; Score 867; DB 14; Length 4334;		
Best Local Similarity 100.0%; Pred. No. 2e-205;		
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1475 TGCTTAGGAACGAGTATTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTCTTAT	1534
Db	2300 TGCTTAGGAACGAGTATTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTCTTAT	2359
Qy	1535 TTAGAAAAGCAATCTAAATATTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	1594
Db	2360 TTAGAAAAGCAATCTAAATATTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	2419
Qy	1595 TAATGACTAGAGAAGAAAGATGAAGATGTTTCATGAAATTAAGAAAGATATTGGATA	1654
Db	2420 TAATGACTAGAGAAGAAAGATGAAGATGTTTCATGAAATTAAGAAAGATATTGGATA	2479
Qy	1655 AATATGGGGATGATCTTAAGGCTATTGGTTTATGGCTCTCTGGTCTGAGACTGATG	1714
Db	2480 AATATGGGGATGATCTTAAGGCTATTGGTTTATGGCTCTCTGGTCTGAGACTGATG	2539
Qy	1715 GGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCAACAGAGGAAGCAGAGTTTCAGCC	1774
Db	2540 GGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCAACAGAGGAAGCAGAGTTTCAGCC	2599
Qy	1775 ATGAATGACAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGTATGAGCAAGATTTCTAC	1834
Db	2600 ATGAATGACAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGTATGAGCAAGATTTCTAC	2659
Qy	1835 TAGATTATGCATCTCAGGTGGAATCAGATTGCGCGCTTACATGCTCAATTTTCTCTA	1894
Db	2660 TAGATTATGCATCTCAGGTGGAATCAGATTGCGCGCTTACATGCTCAATTTTCTCTA	2719
Qy	1895 TTTTCCCGATTATGATTTCAGGTGATATCTTAGAGAAAAGTGTATCAAACTGCTAAATCGG	1954
Db	2720 TTTTCCCGATTATGATTTCAGGTGATATCTTAGAGAAAAGTGTATCAAACTGCTAAATCGG	2779
Qy	1955 TAGAAGCCCAACGTTTCCAGTGCAGTATTTGTCCTTATCGTAGAAGAGCTGTTTGAAT	2014
Db	2780 TAGAAGCCCAACGTTTCCAGTGCAGTATTTGTCCTTATCGTAGAAGAGCTGTTTGAAT	2839
Qy	2015 ATGCAAGCAATGGCGTAATATTCGTGCAAGGACCGACAACATTTCTACCATCTTGA	2074
Db	2840 ATGCAAGCAATGGCGTAATATTCGTGCAAGGACCGACAACATTTCTACCATCTTGA	2899
Qy	2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA	2134

Db	2900 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA	2959
Qy	2135 CGACGACGGCTTCGGTCTTAACGAGCTTAAGCAATCAGATCTTCCTTCAGGTTATG	2194
Db	2960 CGACGACGGCTTCGGTCTTAACGAGCTTAAGCAATCAGATCTTCCTTCAGGTTATG	3019
Qy	2195 ACCATCTGTGCCAGTTCGTAAATGTCGTCTCAACTTTCCGACTCTCGAGAACTTCTTGAAT	2254
Db	3020 ACCATCTGTGCCAGTTCGTAAATGTCGTCTCAACTTTCCGACTCTCGAGAACTTCTTGAAT	3079
Qy	2255 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACAGATATATAGTGG	2314
Db	3080 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACAGATATATAGTGG	3139
Qy	2315 ATGTGTCAAAACGCATACCATTTTGAA	2341
Db	3140 ATGTGTCAAAACGCATACCATTTTGAA	3166
RESULT 10		
AAF29645/C		
ID	AAF29645 standard; DNA; 5142 BP.	
XX	AC	
XX	AAF29645;	
DT	10-APR-2001 (first entry)	
XX		
DE	Plasmid pL2mTFFiv1.	
XX		
KW	Mouse TFF1; mTFF; trefoil protein; antiinflammatory; gastrointestinal;	
KW	anti-ulcer; peptide therapy; gastrointestinal disease; acute colitis;	
KW	Crohn's disease; ulcerative colitis; plasmid pL2mTFFiv1;	
KW	recombinant vector; ds.	
OS	Synthetic.	
PN	WO200102570-A1.	
XX		
PD	11-JAN-2001.	
XX		
PF	05-JUL-2000; 2000WO-EP006343.	
PR	05-JUL-1999; 99EP-00870143.	
XX	(VLAA-) VLAMS INTERUNIVERSITAIR INST BIOTECHNOG.	
PI	Hans WC, Steidler L, Remaut ER;	
XX	WPI; 2001-138142/14.	
XX		
PT	Recombinant Lactococcus lactis for delivering a trefoil peptide useful	
PT	for treating acute or chronic gastrointestinal inflammatory diseases or	
PT	disorders, e.g. acute or ulcerative colitis, acute flare-ups of Crohn's	
XX	disease.	
PS	Claim 18; Fig 1b; 44pp; English.	
XX		
CC	The present sequence is a recombinant vector which may be used to	
CC	generate a recombinant Lactococcus lactis capable of delivering a trefoil	
CC	peptide in vivo. The recombinant microorganism is useful for	
CC	manufacturing an agent for the delivery of a trefoil peptide to the	
CC	gastrointestinal tract, and for treating gastric or intestinal diseases	
CC	or disorders, or lesions caused by gastric or intestinal diseases or	
CC	disorders. The microorganism may also be used for preparing medicament to	
CC	be used for treating gastric and /or gastrointestinal diseases or	
CC	disorders, acute gastrointestinal inflammatory diseases (e.g., acute	
CC	colitis, acute flare-ups of Crohn's diseases, or ulcerative colitis), and	
CC	chronic and spontaneously recurring diseases of the gastrointestinal	
CC	tract comprising Crohn's disease (enteritis regionalis) and ulcerative	
CC	colitis (colitis ulcerosa). Disease states which can be treated by the	
CC	method or compositions comprising the recombinant microorganism or	
CC	trefoil peptides include disorders of and damage to the alimentary canal,	
CC	including the mouth, oesophagus, stomach and large and small intestine,	

CC as well as for the protection and treatment of tissues that lie outside
CC the alimentary canal

XX Sequence 5142 BP; 1617 A; 954 C; 927 G; 1644 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 5; Length 5142;
Best Local Similarity 100.0%; Pred. No. 2.1e-205;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAACAGGATTATTAATAGCTGAATAAGAACGGTCTCCCAATATTCCTAT 1534

Db 2914 TGCTTAGGAACAGGATTATTAATAGCTGAATAAGAACGGTCTCCCAATATTCCTAT 2855

Qy 1535 TTAGAAAAGCAATCTAAAATTTATCTGAAAAGGAATGAGAATAGTGAATGAGCAATAA 1594

Db 2854 TTAGAAAAGCAATCTAAAATTTATCTGAAAAGGAATGAGAATAGTGAATGAGCAATAA 2795

Qy 1595 TAATGACTAGAGAAAGAAATGAGATTGTCATGAATTAAGAAAGCAATATTCGATA 1654

Db 2794 TAATGACTAGAGAAAGAAATGAGATTGTCATGAATTAAGAAAGCAATATTCGATA 2735

Qy 1655 AATATGGGATGATGTTAAGCTATTGGTGTATGGCTCTCTGGCTCAGACTGATG 1714

Db 2734 AATATGGGATGATGTTAAGCTATTGGTGTATGGCTCTCTGGCTCAGACTGATG 2675

Qy 1715 GGCCCTATTCCGATATTGAGATGATGTGTCATGTCAACAGAGAAAGCAGATTCAGCC 1774

Db 2674 GGCCCTATTCCGATATTGAGATGATGTGTCATGTCAACAGAGAAAGCAGATTCAGCC 2615

Qy 1775 ATGAATGCAACCGGTGAGTGGAAAGTGGAGTGAATTTTGATAGCGAGAGATTCAC 1834

Db 2614 ATGAATGCAACCGGTGAGTGGAAAGTGGAGTGAATTTTGATAGCGAGAGATTCAC 2555

Qy 1835 TAGATTATGCAATCAGTGGAAATCAGATTGGCGCTTACACATGCTCAATTTTCTCTA 1894

Db 2554 TAGATTATGCAATCAGTGGAAATCAGATTGGCGCTTACACATGCTCAATTTTCTCTA 2495

Qy 1895 TTTTGGCGATTTATGATTAGTGGGATATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954

Db 2494 TTTTGGCGATTTATGATTAGTGGGATATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 2435

Qy 1955 TAGAAGCCCAACGTTCCAGATCGGATTTGGCCCTTATCGTAGAGAGCTGTTGAAT 2014

Db 2434 TAGAAGCCCAACGTTCCAGATCGGATTTGGCCCTTATCGTAGAGAGCTGTTGAAT 2375

Qy 2015 ATGCAGCAATGCGCTAAATATTCGTGCAAGCAGCAACATTTCTACCATCCTTGA 2074

Db 2374 ATGCAGCAATGCGCTAAATATTCGTGCAAGCAGCAACATTTCTACCATCCTTGA 2315

Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCATCGCATCTGTTATA 2134

Db 2314 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCATCGCATCTGTTATA 2255

Qy 2135 CGACGAGGCTTCGCTTAACTGAAGCAGTAAAGCAATCAGATCTTCCTTCAGGTTATG 2194

Db 2254 CGACGAGGCTTCGCTTAACTGAAGCAGTAAAGCAATCAGATCTTCCTTCAGGTTATG 2195

Qy 2195 ACCATCTGCGAGTCTGTAATGCTGCTCAACTTTCCGACTCTGAGAACTTCGGAAT 2254

Db 2194 ACCATCTGCGAGTCTGTAATGCTGCTCAACTTTCCGACTCTGAGAACTTCGGAAT 2135

Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGCAACGACGGATATATAGTG 2314

Db 2134 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGCAACGACGGATATATAGTG 2075

Qy 2315 ATGTGCAAAACGCAATACCAATTTTGAA 2341

Db 2074 ATGTGCAAAACGCAATACCAATTTTGAA 2048

RESULT 11

AAS17998

ID AAS17998 standard; DNA; 5793 BP.

XX

AC

XX

DT

XX

DE

XX

XX

KW

KW

XX

OS

OS

OS

OS

XX

PN

XX

PD

XX

XX

PP

XX

PR

PR

XX

PA

XX

PI

XX

DR

XX

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

XX

AC

XX

DT

XX

DE

XX

XX

KW

KW

XX

OS

OS

OS

OS

XX

PN

XX

PD

XX

XX

PP

XX

PR

PR

XX

PA

XX

PI

XX

DR

XX

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

XX

AC

XX

DT

XX

DE

XX

XX

KW

KW

XX

OS

OS

OS

OS

XX

PN

XX

PD

XX

XX

PP

XX

PR

PR

XX

PA

XX

PI

XX

DR

XX

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

XX

AC

XX

DT

XX

DE

XX

XX

KW

KW

XX

OS

OS

OS

OS

XX

PN

XX

PD

XX

XX

PP

XX

PR

PR

XX

PA

XX

PI

XX

DR

XX

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

XX

AC

XX

DT

XX

DE

XX

XX

KW

KW

XX

OS

OS

OS

OS

XX

PN

XX

PD

XX

XX

PP

XX

PR

PR

XX

PA

XX

PI

XX

DR

XX

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

XX

AC

XX

DT

XX

DE

XX

XX

KW

KW

XX

OS

OS

OS

OS

XX

PN

XX

PD

XX

XX

PP

XX

PR

PR

XX

PA

XX

PI

XX

DR

XX

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

XX

AC

XX

DT

XX

DE

XX

XX

KW

KW

XX

OS

OS

OS

OS

XX

PN

XX

PD

XX

XX

PP

XX

PR

PR

Qy	1535	TTAGAAAGCAAATCTTAAATTTATCTGAAAGGGAATGAGAAATAGTGAATGGAACAATAA	1594
Db	1028	TTAGAAAGCAAATCTTAAATTTATCTGAAAGGGAATGAGAAATAGTGAATGGAACAATAA	1087
Qy	1595	TAATGACTAGAGAAGAAGAAATGAAGATTGTTCTATGAATTTAAGAAACGAATATTGGATA	1654
Db	1088	TAATGACTAGAGAAGAAGAAATGAAGATTGTTCTATGAATTTAAGAAACGAATATTGGATA	1147
Qy	1655	AATATGGGATGATGTTTAAGCTATTGTTGTTTATGGCTCTCTTGGTGGTCAGACTGATG	1714
Db	1148	AATATGGGATGATGTTTAAGCTATTGTTGTTTATGGCTCTCTTGGTGGTCAGACTGATG	1207
Qy	1715	GGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAGGAAGCAGATTTCAGCC	1774
Db	1208	GGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAGGAAGCAGATTTCAGCC	1267
Qy	1775	ATGAATGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCGAAGAGATTCTAC	1834
Db	1268	ATGAATGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCGAAGAGATTCTAC	1327
Qy	1835	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1894
Db	1328	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1387
Qy	1895	TTTTGCCGATTTATGATTTCAGGTGATATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG	1954
Db	1398	TTTTGCCGATTTATGATTTCAGGTGATATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG	1447
Qy	1955	TAGAAGCCAAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAT	2014
Db	1448	TAGAAGCCAAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAT	1507
Qy	2015	ATGACGGCAATGCGGTAAATTTGTTGTCGAAGGACGCAACATTTCTACCATCCTTGA	2074
Db	1508	ATGACGGCAATGCGGTAAATTTGTTGTCGAAGGACGCAACATTTCTACCATCCTTGA	1567
Qy	2075	CTGTACAGTAGCAATGCGAGTGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA	2134
Db	1568	CTGTACAGTAGCAATGCGAGTGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA	1627
Qy	2135	CGACAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG	2194
Db	1628	CGACAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG	1687
Qy	2195	ACCATCTGTGCGAGTTCGTATGTTCTGTCACACTTTCGACTCTGAGAACTTCGGAT	2254
Db	1688	ACCATCTGTGCGAGTTCGTATGTTCTGTCACACTTTCGACTCTGAGAACTTCGGAT	1747
Qy	2255	CGCTAGAGAAATTTCTGGAAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTG	2314
Db	1748	CGCTAGAGAAATTTCTGGAAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTG	1807
Qy	2315	ATGTGTCAAAACGCATACCAATTTTGA	2341
Db	1808	ATGTGTCAAAACGCATACCAATTTTGA	1834

RESULT 12

ADL16337
ID ADL16337 standard, DNA; 5865 BP.

XX AC ADL16337;

XX AC ADL16337;

DT 06-MAY-2004 (first entry)

XX pBPII backbone vector DNA.

XX Immunogenic; vaccine; lethal infection; protective antigen; PA;

KW lethal factor; LF; anthrax; antibacterial; ds.

XX Unidentified.

OS

PN	US2004028695-A1.
XX	12-FEB-2004.
XX	28-MAR-2003; 2003US-00402466.
PF	12-APR-2002; 2002US-0372152P.
PR	(PARK/) PARK S.
XX	(GIRI/) GIRI L.
PA	Park S, Giril L;
XX	WPI; 2004-168865/16.
DR	Immunogenic composition for preparing a vaccine against a lethal
XX	infection of Bacillus anthracis in an animal, comprises an immunizing
PT	amount of a recombinant B. anthracis protective antigen protein and/or
PT	lethal factor protein.
PS	Disclosure; SEQ ID NO 2; 143pp; English.
XX	The present invention relates to immunogenic compositions useful for
CC	preparing a vaccine against a lethal infection of Bacillus anthracis in
CC	an animal. The compositions comprise an immunising amount of a
CC	recombinant B. anthracis protective antigen (rPA) protein and a
CC	lethal factor (LF) protein. The composition and
CC	methods are useful in protecting against anthrax or lethal infections
CC	caused by B. anthracis. The present sequence is pBPII backbone vector DNA
CC	used in the invention.
XX	Sequence 5865 BP; 1877 A; 1014 C; 1345 G; 1629 T; 0 U; 0 Other;
SQ	

Query Match 27.5%; Score 867; DB 12; Length 5865;
Best Local Similarity 100.0%; Pred. No. 2.2e-205;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1475	TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAAACGGTCTCTCCAAATATTCCTAT	1534
Db	4390	TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAAACGGTCTCTCCAAATATTCCTAT	4449
Qy	1535	TTAGAAAAGCAATCTTAAATTTATCTGAAAAGGAATGAGAAATAGTGAATGGACCAATAA	1594
Db	4450	TTAGAAAAGCAATCTTAAATTTATCTGAAAAGGAATGAGAAATAGTGAATGGACCAATAA	4509
Qy	1595	TAATGACTAGAGAAGAAATGAAGATTCTCATGAAATTAAGGAACGAATATTGGATA	1654
Db	4510	TAATGACTAGAGAAGAAATGAAGATTCTCATGAAATTAAGGAACGAATATTGGATA	4569
Qy	1655	AATATGGGATGATGTTAAGGCTATTGGTGTGTTTATGGCTCTCTTGGTGGTCAGACTGATG	1714
Db	4570	AATATGGGATGATGTTAAGGCTATTGGTGTGTTTATGGCTCTCTTGGTGGTCAGACTGATG	4629
Qy	1715	GGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAGGAAGCAGATTTCAGCC	1774
Db	4630	GGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAGGAAGCAGATTTCAGCC	4689
Qy	1775	ATGAATGGACAACCGGTGAGTGGAAAGTGAATTTTGTATAGCGAAGAGATTCTTAC	1834
Db	4690	ATGAATGGACAACCGGTGAGTGGAAAGTGAATTTTGTATAGCGAAGAGATTCTTAC	4749
Qy	1835	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1894
Db	4750	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	4809
Qy	1895	TTTTGCCGATTTATGATTTCAGGTGGATACTTAGAGAAAGTGTATCAAACTGCTAAATCGG	1954
Db	4810	TTTTGCCGATTTATGATTTCAGGTGGATACTTAGAGAAAGTGTATCAAACTGCTAAATCGG	4869
Qy	1955	TAGAAGCCAAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAT	2014
Db	4870	TAGAAGCCAAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAT	4929

QY 2015 ATGAGGCAATCGCGTAATATTCGTGTGCAAGGACCGAACATTTCTACCAATCTTGA 2074
Db |||||
QY 4930 ATGAGGCAATCGCGTAATATTCGTGTGCAAGGACCGAACATTTCTACCAATCTTGA 4989
Db |||||
QY 2075 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGCTGCGCATCATCGCATCTGTATA 2134
Db |||||
QY 4990 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGCTGCGCATCATCGCATCTGTATA 5049
Db |||||
QY 2135 CGACGAGCGTTCGGTCTTAATCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db |||||
QY 5050 CGACGAGCGTTCGGTCTTAATCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 5109
Db |||||
QY 2195 ACCATCTGCGGATTCGTAAATGTCGTGCAACTTTCCGACTCTGAGAACTTCGGAAT 2254
Db |||||
QY 5110 ACCATCTGCGGATTCGTAAATGTCGTGCAACTTTCCGACTCTGAGAACTTCGGAAT 5169
Db |||||
QY 2255 CGCTAGAGATTTCTGGAATGCGATTCAGGATGCGACGACGACGATATATAGTGG 2314
Db |||||
QY 5170 CGCTAGAGATTTCTGGAATGCGATTCAGGATGCGACGACGACGATATATAGTGG 5229
Db |||||
QY 2315 ATGTGTCAAAACGCATACCAATTTTGAA 2341
Db |||||
QY 5230 ATGTGTCAAAACGCATACCAATTTTGAA 5256
Db |||||

RESULT 13

AA517997
ID AA517997 standard; DNA; 5943 BP.
XX
AC AA517997;
XX
DT 12-MAR-2002 (first entry)
XX
DE Plasmid pMOL1807.
XX
XX pMOL1807; cyclic; circular; GalE; ds; UDP-galactose; UDP-glucose;
KW amplification unit copy number; industrial protein production;
KW B. subtilis; dal locus; MOL1875.
XX
OS Synthetic.
OS Bacillus licheniformis.
OS Bacillus subtilis.
OS Staphylococcus aureus.
OS Chimeric.
XX
XX WO200190393-A1.
XX
XX 29-NOV-2001.
XX
XX 23-MAY-2001; 2001WO-DK000356.
XX
XX 24-MAY-2000; 2000DK-00000824.
PR 30-MAY-2000; 2000US-0208052P.
XX
XX (NOVO) NOVOZYMES AS.
XX
XX Rasmussen MD;
XX
XX WPI; 2002-097661/13.
DR
XX
XX Increasing copy number of an amplification unit in a host cell
PT chromosome, for industrial production of polypeptides, comprises
PT introducing an expression cassette having a gene of interest and a gene
PT non-functional in the chromosome.
XX
XX Example; Page 63-64; 69pp; English.
PS
XX
XX The invention relates to increasing the number of copies of an
CC amplification unit integrated into a host cell chromosome, comprising
CC introducing a nucleic acid construct consisting of an amplification unit
CC having an expression cassette with a copy of a gene of interest, and an
CC expressible copy of the non-functional chromosomal gene (where the host
CC cell has the non-functional gene), where a copy of the amplification unit

CC integrates into the host cell chromosome, the host cell becomes
CC susceptible to an inhibitory compound endogenously produced by the host
CC cell when cultivated in a medium comprising a precursor, and at least one
CC copy of the amplification unit integrates into host cell chromosome,
CC cultivating the host cell in a medium comprising the precursor, where a
CC chromosomally integrated copy of the amplification unit is duplicated or
CC multiplied on the host cell chromosome, selecting a host cell comprising
CC two or more chromosomally integrated copies of the amplification unit and
CC optionally performing one or more cycles of using the host cell selected
CC in each new cycle, where the number of chromosomally integrated copies
CC of the amplification unit increases with each repeat. The method is
CC useful for increasing the copy number of a chromosomally integrated
CC expression cassette in a microbial strain devoid of antibiotic resistance
CC markers for use in industrial production of polypeptides in high yield.
CC The present sequence is replication deficient plasmid used to integrate a
CC copy of the GalE gene from B. subtilis into the dal locus of strain
CC MOL1875 which lacks a functional GalE gene. The GalE gene converts toxic
CC UDP-galactose to non-toxic UDP-glucose
XX
SQ Sequence 5943 BP; 1839 A; 1023 C; 1459 G; 1622 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 6; Length 5943;
Best Local Similarity 100.0%; Pred. No. 2.2e-205;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAAGCAGGTTATTAATAGCTGAATAGAACGGTCTCTCCAAATATTCCTAT 1534
Db |||||
QY 968 TGCTTAGGAAGCAGGTTATTAATAGCTGAATAGAACGGTCTCTCCAAATATTCCTAT 1027
Db |||||
QY 1535 TTAGAAAAGCAATCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATA 1594
Db |||||
QY 1028 TTAGAAAAGCAATCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATA 1087
Db |||||
QY 1595 TAATGACTAGAGAAGAAAGAAATGAAGATTTCTCATGAAATTAAGGAACGAATTTGGATA 1654
Db |||||
QY 1088 TAATGACTAGAGAAGAAAGAAATGAAGATTTCTCATGAAATTAAGGAACGAATTTGGATA 1147
Db |||||
QY 1655 AATATCGGGATGATGTTAAAGGCTATTGGTGTGTTTATGGCTCTCTTGGTCTGAGACTGATG 1714
Db |||||
QY 1148 AATATCGGGATGATGTTAAAGGCTATTGGTGTGTTTATGGCTCTCTTGGTCTGAGACTGATG 1207
Db |||||
QY 1715 GGCCCTATTTCGGATTTAGAGATGATGTGTGTCATGTCAACAGAGGAGAGAGTTGAGCC 1774
Db |||||
QY 1208 GGCCCTATTTCGGATTTAGAGATGATGTGTGTCATGTCAACAGAGGAGAGAGTTGAGCC 1267
Db |||||
QY 1775 ATGAATGGACAACCGGTGAGTGAAGGTGAAGTGAATTTTGATAGCGAAGAGATTTCTAC 1834
Db |||||
QY 1268 ATGAATGGACAACCGGTGAGTGAAGGTGAAGTGAATTTTGATAGCGAAGAGATTTCTAC 1327
Db |||||
QY 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
Db |||||
QY 1328 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1387
Db |||||
QY 1895 TTTTGGCGATTATGATTCAGGTGGATCTATAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db |||||
QY 1388 TTTTGGCGATTATGATTCAGGTGGATCTATAGAGAAAGTGTATCAAACTGCTAAATCGG 1447
Db |||||
QY 1955 TAGAAGCCCAAAACGTTCCAGATGCGATTGTGCCCCCTTATCGTAGAAGAGCTGTTTGAAT 2014
Db |||||
QY 1448 TAGAAGCCCAAAACGTTCCAGATGCGATTGTGCCCCCTTATCGTAGAAGAGCTGTTTGAAT 1507
Db |||||
QY 2015 ATGACGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGAACCAATTTCTACCATCTTGA 2074
Db |||||
QY 1508 ATGACGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGAACCAATTTCTACCATCTTGA 1567
Db |||||
QY 2075 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGCTGCGCATCATCGCATCTGTATA 2134
Db |||||
QY 1568 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGCTGCGCATCATCGCATCTGTATA 1627
Db |||||
QY 2135 CGACGAGCGTTCGGTCTTAATCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db |||||
QY 1628 CGACGAGCGTTCGGTCTTAATCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 1687
Db |||||

Qy 2195 ACCATCTGTGCAGTTCGTAATGTCTGGTCAACTTTCGACTCTGAGAAACTTCTGGAAT 2254
Db |||||||
Db 1688 ACCATCTGTGCAGTTCGTAATGTCTGGTCAACTTTCGACTCTGAGAAACTTCTGGAAT 1747
Qy 2255 CGCTAGAGAAATTTCTGGAATCGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314
Db |||||||
Db 1748 CGCTAGAGAAATTTCTGGAATCGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 1807
Qy 2315 ATGTGTCAAAACGCATACCAATTTGAA 2341
Db |||||||
Db 1808 ATGTGTCAAAACGCATACCAATTTGAA 1834

RESULT 14

AAD49392/C
ID AAD49392 standard; DNA; 6661 BP.

XX AC
AC AAD49392;

XX 07-MAR-2003 (first entry)

XX pMOL995 plasmid DNA.

XX Pectate lyase; EC 4.2.2.2; detergent; surfactant; cleaning; dishwashing;
KW fabric stain removal; fabric whiteness maintenance; fabric softening;
KW fabric colour appearance; fabric dye transfer inhibition; de-pilling;
KW dental application; oral application; colour clarification; enzyme;
KW clay stain removal; ds.

XX OS
OS Unidentified.

XX Key Location/Qualifiers

PH misc_feature

FT 1..1962 a

FT /note= "Encodes PUB110 plasmid"

FT 1963..2305

FT /tag= b

FT /note= "Encodes transcriptional terminator from amyL gene
of B. licheniformis ATCC14580 and a few introduced

FT restriction sites"

FT 2306..3766

FT /tag= c

FT /note= "Encodes mature part of alpha-amylase"

FT /tag= d

FT /note= "Encodes the promoter and signalpeptide of alpha-

FT amylase"

FT 4076..6661

FT /tag= e

FT /note= "Encodes PUB110 plasmid"

XX WO200292741-A2.

PN 21-NOV-2002.

XX 14-MAY-2002; 2002WO-DK000315.

XX 14-MAY-2001; 2001DK-00000755.

PR 14-MAY-2001; 2001US-0290738P.

XX (NOVO) NOVOZYMES AS.

XX Eskelund MB, Schuelein M, Nielsen VS, Smets J;

XX WPI; 2003-059005/05.

XX New polypeptide useful as ingredient of detergent composition for
PT cleaning fabric, dishware or hard surface, encoded by DNA sequence
PT endogenous to strain of Bacillus subtilis, comprises pectate lyase
PT activity.

XX Disclosure; Page 84-87; 95pp; English.

XX

CC The present invention relates to pectate lyase (EC 4.2.2.2) proteins and
CC polynucleotides encoding such proteins. The invention also relates to
CC detergent compositions comprising a surfactant and pectate lyase enzyme.
CC The detergent is useful for cleaning a fabric, a dishware or hard surface
CC to provide superior cleaning performance, for fabric cleaning and/or
CC fabric stain removal and/or fabric whiteness maintenance and/or fabric
CC softening and/or fabric colour appearance and/or fabric dye transfer
CC inhibition, for cleaning hard surfaces such as floors, walls or bathroom
CC tiles, for hand and machine dishwashing and for oral and/or dental
CC applications. They are useful for colour clarification, de-pilling and in
CC clay stain removal. The present sequence is pMOL995 plasmid DNA
XX

SQ Sequence 6661 BP; 1859 A; 1534 C; 991 G; 2277 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 10; Length 6661;

Best Local Similarity 100.0%; Pred. No. 2.3e-205;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGACGAGTTTATTAATAGCTGAATAGAACGGTGTCTCCAAATATTTCTTAT 1534

Db |||||||

Qy 1535 TTAGAAAAGCAATCTAAAAATTATCTGAAAAGGGAATGAGAATAGTGAGCAATAA 1594

Db |||||||

Qy 5120 TTAGAAAAGCAATCTAAAAATTATCTGAAAAGGGAATGAGAATAGTGAGCAATAA 5061

Qy 1595 TAATGACTAGAGAAGAAGAAATGAAGATTGTTCTATGAAATTAAGGAACGAATATTGGATA 1654

Db |||||||

Qy 5060 TAAATGACTAGAGAAGAAGAAATGAAGATTGTTCTATGAAATTAAGGAACGAATATTGGATA 5001

Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGCTCTCTTGGTCTGTCAGACTGATG 1714

Db |||||||

Qy 5000 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGCTCTCTTGGTCTGTCAGACTGATG 4941

Qy 1715 GGCCCTATTCGGATATTGAGATGATGTGTCTATGTCAACAGAGAAAGAGAGTTACGCC 1774

Db |||||||

Qy 4940 GGCCCTATTCGGATATTGAGATGATGTGTCTATGTCAACAGAGAAAGAGAGTTACGCC 4881

Qy 1775 ATGAATGGACAACCGGTGAGTGGAAAGTGAAGTGAATTTTGTATAGCGAAGAGATTTCTAC 1834

Db |||||||

Qy 4880 ATGAATGGACAACCGGTGAGTGGAAAGTGAAGTGAATTTTGTATAGCGAAGAGATTTCTAC 4821

Qy 1835 TAGATATTGATCTCAGGTGGAATCAGATTGGCGCTTACACATGGTCAATTTTCTCTA 1894

Db |||||||

Qy 4820 TAGATATTGATCTCAGGTGGAATCAGATTGGCGCTTACACATGGTCAATTTTCTCTA 4761

Qy 1895 TTTTGGCGATTTATGATTCAGGTGGATATCTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954

Db |||||||

Qy 4760 TTTTGGCGATTTATGATTCAGGTGGATATCTAGAGAAAGTGTATCAAACTGCTAAATCGG 4701

Qy 1955 TAGAAGCCCAACGTTTCCAGATGCGATTTGTGCCCTTATCGTAGAGAGCTGTTGAAAT 2014

Db |||||||

Qy 4700 TAGAAGCCCAACGTTTCCAGATGCGATTTGTGCCCTTATCGTAGAGAGCTGTTGAAAT 4641

Qy 2015 ATGACGGCAATGCGGTAAATATTCTGTGTGCAAGCGACGCAACAACTTCTACCATCTTGA 2074

Db |||||||

Qy 4640 ATGACGGCAATGCGGTAAATATTCTGTGTGCAAGCGACGCAACAACTTCTACCATCTTGA 4581

Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCAATCATCGCATCTGTTATA 2134

Db |||||||

Qy 4580 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCAATCATCGCATCTGTTATA 4521

Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194

Db |||||||

Qy 4520 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 4461

Qy 2195 ACCATCTGTCCAGTTCGTAATGTCTGGTCAACTTTTCCGACTCTGAGAACTCTCTGGAAT 2254

Db |||||||

Qy 4460 ACCATCTGTCCAGTTCGTAATGTCTGGTCAACTTTTCCGACTCTGAGAACTCTCTGGAAT 4401

Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACACGGATATATAGTGG 2314

Db |||||||

Qy 4400 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACACGGATATATAGTGG 4341

XX

QY 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341
ADL16336
Db 4340 ATGTGTCAAAACGCATACCATTTTGAA 4314

RESULT 15
ADL16336
ID ADL16336 standard; DNA; 6694 BP.
XX
AC ADL16336;
DT 06-MAY-2004 (first entry)
XX
XX pBPI backbone vector DNA.
XX
XX Immunogenic; vaccine; lethal infection; protective antigen; PA;
KW lethal factor; LF; anthrax; antibacterial; ds.
XX
XX Unidentified.
XX
XX US2004028695-A1.
XX
XX 12-FEB-2004.
XX
XX 28-MAR-2003; 2003US-00402466.
XX
XX 12-APR-2002; 2002US-0372152P.
XX
XX (PARK/) PARK S.
PA (GIRI/) GIRI L.
XX
XX Park S, Giri L;
PI
XX WPI; 2004-168865/16.
XX
PT Immunogenic composition for preparing a vaccine against a lethal
PT infection of Bacillus anthracis in an animal, comprises an immunizing
PT amount of a recombinant B. anthracis protective antigen protein and/or
PT lethal factor protein.
XX
XX Disclosure; SEQ ID NO 1; 143pp; English.
XX
XX The present invention relates to immunogenic compositions useful for
XX preparing a vaccine against a lethal infection of Bacillus anthracis in
XX an animal. The compositions comprise an immunising amount of a
XX recombinant B. anthracis protective antigen (rPA) protein and a
XX recombinant B. anthracis lethal factor (rLF) protein. The composition and
XX methods are useful in protecting against anthrax or lethal infections
XX caused by B. anthracis. The present sequence is pBPI backbone vector DNA
XX used in the invention.
XX
SQ Sequence 6694 BP; 2086 A; 1229 C; 1544 G; 1835 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 12; Length 6694;
Best Local Similarity 100.0%; Pred. No. 2.3e-205;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAAGACAGTATTATAGCTGAATGAAGACGGTGCTCTCCAAATATTTCTAT 1534
Db 5219 TGCTTAGGAAGACAGTATTATAGCTGAATGAAGACGGTGCTCTCCAAATATTTCTAT 5278

QY 1535 TTAGAAAAGCAATCTAAAATTTCTGAAAAGGAATGAGATAGTGAATGGACCAATAA 1594
Db 5279 TTAGAAAAGCAATCTAAAATTTCTGAAAAGGAATGAGATAGTGAATGGACCAATAA 5338

QY 1595 TAATGACTAGAGAAGAAGATGAAGATTTTCATGAAATTAAGGAACGAATATTGGATA 1654
Db 5339 TAATGACTAGAGAAGAAGATGAAGATTTTCATGAAATTAAGGAACGAATATTGGATA 5398

QY 1655 AATATGGGGATGATGTTAAGGCTATTGGTGTATGGCTCTCTTGGTCGTCAGACTGATG 1714
Db 5399 AATATGGGGATGATGTTAAGGCTATTGGTGTATGGCTCTCTTGGTCGTCAGACTGATG 5458

QY 1715 GGCCCTATTCCGATATTGAGATGATGTCTCATGTCAACAGAGGAAGCAGAGTTCCAGCC 1774
Db 5459 GGCCCTATTCCGATATTGAGATGATGTCTCATGTCAACAGAGGAAGCAGAGTTCCAGCC 5518

QY 1775 ATGAATGGACAACCCGGTGAAGAGTGGAAGTGAATTTTGTATAGCAAGAGATTCTTAC 1834
Db 5519 ATGAATGGACAACCCGGTGAAGAGTGGAAGTGAATTTTGTATAGCAAGAGATTCTTAC 5578

QY 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
Db 5579 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 5638

QY 1895 TTTTGGCCGATTATGATTTCAGGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db 5639 TTTTGGCCGATTATGATTTCAGGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 5698

QY 1955 TAGAAGCCCAAAAGCTTCCACGATGCGATTTGTGCCCCATTATCGTAGAAGAGCTGTTGAAT 2014
Db 5699 TAGAAGCCCAAAAGCTTCCACGATGCGATTTGTGCCCCATTATCGTAGAAGAGCTGTTGAAT 5758

QY 2015 ATGAGGCAAAATGCGCGTAATATTCTGTGCAAGGACGACAAATTTCTACCATCCTTGA 2074
Db 5759 ATGAGGCAAAATGCGCGTAATATTCTGTGCAAGGACGACAAATTTCTACCATCCTTGA 5818

QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCGATCATCGCATCTGTTATA 2134
Db 5819 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCGATCATCGCATCTGTTATA 5878

QY 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194
Db 5879 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 5938

QY 2195 ACCATCTGTGCCAGTTCTGTAATCTCTGGTCAACTTTCCGACTCTGAGAACTTCTGGAAT 2254
Db 5939 ACCATCTGTGCCAGTTCTGTAATCTCTGGTCAACTTTCCGACTCTGAGAACTTCTGGAAT 5998

QY 2255 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314
Db 5999 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 6058

QY 2315 ATGTGCAAAAACGCATACCATTTTGAA 2341
Db 6059 ATGTGCAAAAACGCATACCATTTTGAA 6085

Search completed: February 11, 2006, 18:11:22
Job time : 1228 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2006, 18:11:35 ; Search time 8017 Seconds
(without alignments)
18418.365 Million cell updates/sec

Title: US-10-681-086-1
Perfect score: 3156
Sequence: 1 ggatccagcgaggtacgagc.....atcggtatgatgcgaattc 3156

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_hic:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_est7:*
- 9: gb_gss1:*
- 10: gb_gss2:*
- 11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	175.2	5.6	1064	10 AY421133	Mus muscu
2	175.2	5.6	1607	4 AK087433	Mus muscu
3	175.2	5.6	1677	4 AK075782	Mus muscu
4	173.6	5.5	1064	10 AY421131	Homo sapi
5	173.6	5.5	2327	4 AK013138	Mus muscu
6	171.4	5.4	1737	4 AK076469	Mus muscu
7	167.2	5.3	1528	4 CR860400	Pongo pyg
8	161.4	5.1	911	5 BX325788	BX325788
9	159.8	5.1	794	6 CA778095	ip1h09.x
10	159.8	5.1	810	6 CA778022	ip1h09.x
11	157.6	5.0	862	8 CX934609	JGI CA03
12	157.4	5.0	805	8 CX438675	JGI XZG67
13	157.4	5.0	1064	10 AY421132	Pan trogl
14	157.2	5.0	787	2 BG752866	602732509
15	157.2	5.0	871	8 DN019683	JGI CAAR2
16	157.2	5.0	901	8 DR871972	JGI CABG1
17	156	4.9	897	8 DN022379	JGI CAAR3
18	155.8	4.9	844	8 DN080099	JGI CABD1
19	155.6	4.9	885	8 DN083438	JGI CABD1
20	155.6	4.9	906	8 CX389898	JGI_XZT37
21	155.4	4.9	609	7 CF900218	A0312H09-
22	155.4	4.9	609	7 CF900995	A0322E10-

23	155.4	4.9	615	7 CF900454	CF900454 A0315G03-
24	155	4.9	677	2 BG928003	BG928003 HNC45-1-H
25	154.4	4.9	1004	2 BB610357	BB610357 BB610357
26	154.2	4.9	895	2 BI251895	BI251895 602952509
27	154	4.9	978	5 BU519273	BU519273 AGENCOURT
28	153.8	4.9	792	7 CV675649	CV675649 i631807.k
29	153	4.8	797	2 BG747688	BG747688 602705159
30	152.6	4.8	580	7 CN347398	CN347398 170005830
31	152.2	4.8	789	8 DN027048	DN027048 JGI CAAR6
32	152.2	4.8	864	8 DN028149	DN028149 JGI CAAR6
33	152.2	4.8	874	8 DN025254	DN025254 JGI CAAR5
34	152.2	4.8	883	7 CR422328	CR422328 CR422328
35	152.2	4.8	891	7 CR567595	CR567595 CR567595
36	152	4.8	908	6 CB566539	CB566539 AGENCOURT
37	151.2	4.8	733	5 BX325548	BX325548 BX325548
38	151.2	4.8	867	8 DN025549	DN025549 JGI CAAR5
39	150.8	4.8	955	5 BU526482	BU526482 AGENCOURT
40	150.6	4.8	781	8 CX327170	CX327170 JGI_XZT15
41	150.6	4.8	786	8 CX969308	CX969308 JGI_CAAP3
42	150.6	4.8	882	8 CX362459	CX362459 JGI_XZT42
43	150.6	4.8	884	8 CX468737	CX468737 JGI_XZG33
44	150.6	4.8	889	8 DN033808	DN033808 JGI CAAR1
45	150.6	4.8	911	8 CX391320	CX391320 JGI_XZT38

ALIGNMENTS

RESULT 1
AY421133
LOCUS 1064 bp DNA linear GSS 17-DEC-2003
DEFINITION Mus musculus GCAT gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY421133
VERSION AY421133.1 GI:39777090
KEYWORDS GSS.
ORGANISM Mus musculus (house mouse)
SOURCE Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1064)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1064)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
source Location/Qualifiers
1..1064
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1064
/gene="GCAT"
/locus_tag="HGM7455"

Query Match 5.6%; Score 175.2; DB 10; Length 1064;
Best Local Similarity 49.9%; Pred. No. 7.5e-35;
Matches 475; Conservative 0; Mismatches 468; Indels 9; Gaps 1;

Qy	158	CARCAATATTATTTAGGCTCGCAAGCGATAGACGTTTGATCGATGAGCCCAACACGAT	217
Db	19	CCAATAACTACTGGGCTGAGAGCCACCTCGAGTGATCCAGGCGGTCTGCAGACTC	78
Qy	218	TGCAGCAATTTGGGACAGGAGCAGCGGTTTACGTTTAAACGACAGCAATTCGCTCTGGC	277
Db	79	TGGAGGAGTTTGGAGCTGGAATCTAGTTTCTAGTTTATCTGTGGACCCAGAGCATCC	138
Qy	278	ATGAAAGCTAGAAAAGAGATTGCGAGCTTTAACTGACAGAGCGGCCCTCTGTTTT	337
Db	139	ATAAGAATCTAGAAGCCAGATAGCCCACTTCCACGAGCTGAGGACGCCATCCTCTATC	198
Qy	338	CGAGGGTTACTTGGCCAAATGCTGGTGTCTTTCATCTTCCGCAAGAAAGCAAGATGCA	397
Db	199	CCAGCTGTTTCGATGCGCAACCTGCTGCTATTTAGGCTTCTGCTACCCAGAGATCGG	258
Qy	398	TTTTAAAGTGACCAAGCTCAATCATGCAAGATGATGACGCGCTGCGACTTTCTAAGGCTG	457
Db	259	TTCTGTGATGAGCTGNAACCATGCTTCCATCATTTAGCGGCATCGACTGTGCAAGGCC	318
Qy	458	ATACAGTTGTTTTCGGCATATTGATATGAATGATCTTTGAAAACAAGCTGAATGAACAC	517
Db	319	ACAAAGTACCGTTAGCCGCCACCTGGACATGGCTGACCTGGAAGCCAAAGCTGAAGGAGCTC	378
Qy	518	AGCGTTATCAGCGCGGTTTATCTGTAACAGACGGAGTATTCAGCATGGATGGCAACATCG	577
Db	379	AGAAAGCAGGCTGCGGCTGTGTGGCCACAGATGAGCTTCTCCATGATGGTGACATCG	438
Qy	578	CCCTCTTTGATCAGATCATCTCACTTGCGAAACGCTATCATGCGCTTCTGTCGCTGTGATG	637
Db	439	CTCCCTACAGGACATCTGCGGCTTGCGGCCCAATATGTTGGTGCCTGGTCTTTGTGATG	498
Qy	638	ATGCCACGCAACAGGAGTTTGGGCGATTTCGGACAAAGAAACGAGTGAATCTTTGGTG	697
Db	499	AATGCCATGCCATGGCTTCTTGGGCCCAACAGACGGGGCACAGATGAGCTGTAGTGTG	558
Qy	698	TTTGTCCGACATGTTATCGGCACCTTAAGCAAG-----CTGTTGGCGGGAG	748
Db	559	TCATGGACAGGTCAACATATCACTCCACCTTGGGGAAGCGCTGGTGGAGCATCAG	618
Qy	749	GAGCTTTTGGCGGAGGATCAGCGTCTTTCATCGACTTTTTCCTGAAACCATGCCAGAACAT	808
Db	619	GGGATATACACCGGGCTGAGCCACTGGTGTCCCTGCTACGGCAGGTTCTCGGCCCT	678
Qy	809	TTATCTTTCAACCGCTATTTCGGCAGCGACTGTGGGCTGCTCAGAGGCTTTCAACA	868
Db	679	ACCTCTTCTCAACAGCGCTGCGCCCTGCTGTAGTCGGCTGTCTTCTAAGGCGCCTGGACC	738
Qy	869	TCATTGAAGCAGCAGGGAACGACGCTTTTATTTCTTATATCAGCATGATCAGAA	928
Db	739	TGCTCATGGAGAGCAATGCCATCATCCAGTCTATGGCTGCCAAGACCCGGCGTTCCGCA	798
Qy	929	CCAGTCTGAAGATATGGGTTATGTGTGTAAGGAGATCACACCCGATATTCTCTGTAG	988
Db	799	GTAAGATGGAAGCTGCTGTTTACTGTCTCAGAGCGGATCACCCCATCTGCCCCTGTGA	858
Qy	989	TCATTGGCGATGCCCAATAAACCGTCTTATTTGCTGAAAAACTGACGGCAAGGAATTT	1048
Db	859	TGCTGGGAGATGCACGGTGTCTTCTCAATGGCAGATGACATGTTAAAGAAAGGTATCT	918
Qy	1049	ATGCTCTGCCATTCGGCGCCCAACCGTTGCGCGGGTGAAGCGGATTCG	1100
Db	919	TTGTATCGGATTTAGCTACCTGTGGTCCCTAAGGGCAAGGCCCGATCCG	970
RESULT 2			
AK087433			
LOCUS	AK087433	1607 bp	mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus 0 day neonate eyecell cDNA, RIKEN full-length enriched library, clone:R130112P7 product:glycine C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase), full insert sequence.		
ACCESSION	AK087433		

VERSION	AK087433.1	GI:26352483
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 10349636	
AUTHORS	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 11042159	
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 11076861	
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	
AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1607) Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hasegawa, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	
AUTHORS	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers 1. .1607	
COMMENT		
FEATURES		
source		

November 2005

Published Applications Nucleic Acid and Published Applications Amino Acid database searches now generate two sets of results each. The Published Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published Applications New databases; older published applications make up the Published Applications Main databases.

Searches run against Nucleic Acid Published Applications produce two sets of results, with the extensions **npbm** (Published Applications NA_Main) and **rnpbm** (Published Applications NA_New).

Searches run against Amino Acid Published Applications produce two sets of results, with the extensions **rapbm** (Published Applications AA_Main) and **rapbn** (Published Applications AA_New).

This Page Blank (uspto)

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:EI130112F17"
/db_xref="taxon:10090"
/clone="EI130112F17"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
20. 1270
CDS
/notes="unnamed protein product; glycine
C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A
ligase) (MGI:1349389)
putative"
/codon_start=1
/protein_id="BAC39872.1"
/db_xref="GI:26352484"
/translation="MWSFHWGALSPGRHSAALQRLCILDSLEGIRGATWSE
RVITSQFSIRVDGSGGILNFCANNYIGLSHPAVIQGLQTLSEFGAGLSSTFI
CQTQSIRKQLEAKIAHFQREDAIYPSCFDANAGLFEALLPEDAVLSDLNHASII
DGIKLCARHYRHLDMADLEAKLEAKQRLRLVATDGAFCMDGDIAPLODICLA
AOYGLVFDVDECHATFLPTGRGTDELLGVMDQVTIINSTGLKAGSAGGYTQPE
PLVSLARSRPVLFSNLSLPPAVVGCASKALDLMESNAIIQSMAAKTRFRSKMEA
GFTVSGADHPICVPLMLGDRLSSQADMDMLKGI FVIGFSYVVPVPGKARIRVQISAV
HSEBIDRDCVEAFVEVGRHLGALP"
ORIGIN
Query Match 5.6%; Score 175.2; DB 4; Length 1607;
Best Local Similarity 49.9%; Pred. No. 8.4e-35;
Matches 475; Conservative 0; Mismatches 468; Indels 9; Gaps 1;
QY 158 CAACAAATATTATTTAGGCTCGCAAGCGATGACAGCTTTTGATCGATGCAGCCCAACAGCAT 217
DB 225 CCAATAACTACCTGGGCTGAGCAGCACCCCTGCAGTGATCCAGGCGTCTCAGACTC 284
QY 218 TGCAGCAATTTGGACAGGAAGCAGCGGTTACGTTTAAACGACAGCAATTCGGTCTGGC 277
DB 285 TGGAGGAGTTTGGAGCTGGACTCAGTTCTACTCGATTATCTCTGGGACCCAGAGCATCC 344
QY 278 ATGAAAAGCTAGAAAAGAGATTGCCAGCTTTAAACATGCAGAGAGCGGCCCTCGTGTIT 337
DB 345 ATAAGAACTAGAAAGCCAGATAGCCCACTTCCACAGCGTGAGGAGCCCATCTCTATC 404
QY 338 CGAGCGGTACTTTGGCCAAATGTGCGTGTCTCTTTTCATCTTCCGAGAAAAGAGATGTCA 397
DB 405 CCAGCTGTTTCGATGCCAACGCTGGCCTATTATGAGGCTTTTGCTGACCCAGAGATGCGG 464
QY 398 TTTTAAAGTGACAGCTCAATCATGCGAAGTAGATGCGAGCGCTGCCGACTTTCTTAAGCTG 457
DB 465 TTCTGTGAGATGAGCTGAAACCATGCTTCCATCATTTGACGGCATCCGACTGTGCAAGGCC 524
QY 458 ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTGAAGAACAAAGCTGAATGAAACAC 517
DB 525 ACAAGTACCGTTACCGCCACTCGACATGCTGACCTGGGAAGCAAGCTGAAGAGGCTC 584
QY 518 AGCGTTATCAGCGCGGTTTTATCGTAAACAGACGAGGATTTTCAGCATGATGGCACAATCG 577
DB 585 AGAAGCACAGCTCGCGCTGCTGGCCACACAGATGGAGCTTTCTCCATGATGGTGACATCG 644
QY 578 CCCCTCTTGATCATGATCATCTCATCTGCGAAAGCGTATCATGCGCTTCGCTGGTGTGATG 637
DB 645 CTCCCTTACAGGACATCTGCGCGCTTGCCGCCCAATATGTTGTCCTGCTTTTGTGATG 704
QY 638 ATGCCACGCAACAGGAGTTTGGGGATTTGGGGATTTGGGACAAGCAAGGAGTAATCTTTGGTG 697
DB 705 AATGCCATGCCATCGGCTTTCTTGGGCCCAAGACGGGCAAGATGAGCTGCTGAGTG 764
QY 698 TTTGTCCCGCATTTTATCGGCACCTTTAAGCAAG-----CTGTTGGGCGCGAAG 748
DB 765 TCATGGACAGGTCACCATCATCACTCCACCTGGGAGGCGCTGGTGGAGCATCAG 824
QY 749 GAGGTTTGGCGGAGGATCAGCGTCTTCAATCGACTTTTTCGTAACCATGCCAGAACAT 808

```

```

DB 825 GGGGATATACACCGGGCCTGAGCCACTGGTGTCCCTGCTACGGCAGCGTTCTCGGGCCT 884
QY 809 TTATCTTTCAAACCGCTATTTCGCCAGCCAGCTGTGGCTGTCTCAGAGGCTTTCAACA 868
DB 885 ACCTCTTCTCCACAGCCTGCCCTGTCTAGTGGCTGTGCTTCTAAGCCCTGGACC 944
QY 869 TCATTGAAGCAGCAGGGAACACGACAGCTTTTATTTTCTTATATCAGCATGATCAGAA 928
DB 945 TGCTCATGAGAGCAATGCCATCATCCAGTCTATGGTGTGCAAGACCCGCGGTTCGCA 1004
QY 929 CCAGTCTGAAGATATGGGTTATGTGTGAAGAGGATCACACCGATTATTCCTCTAG 988
DB 1005 GTAAGATGAAGCTGCTGGTTTTTCTGTCTCAGAGCCGATCACCCCATCTGCCCTGTA 1064
QY 989 TCATTGGCGATGCCATAAAACGGTCTCTATTTCCTGAAAACTGCAGGGCAAGGAATTT 1048
DB 1065 TGGTGGAGATGCACGGTTGTCTTCTCAATGCGCAGATGACATGTTAAAGAAAGGTATCT 1124
QY 1049 ATGCTCTGCGCATTCGGCGCCCAACCGTTGCGCGCGGTGAAAGCCGGATTGCG 1100
DB 1125 TTGTCTACGATTTAGCTACCTGTGCTCCCTTAAGGGCAAGGCCCGGATCCG 1176

RESULT 3
AK075782 1677 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
DEFINITION enriched library, clone:l810021l23 product:glycine
C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase),
full insert sequence.
ACCESSION AK075782
VERSION AK075782.1 GI:26096539
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
2
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
3
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tachiro, H., Itoh, M.,
Sumi, N., Ihii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861
4
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861
5
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)

```

REFERENCE	6 (bases 1 to 1677)	
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers	
FEATURES	source	
	1..1677	/organism="Mus musculus"
		/mol_type="mRNA"
		/strain="C57BL/6J"
		/db_xref="FANTOM_DB:1810021L23"
		/db_xref="taxon:10090"
		/clone="1810021L23"
		/sex="male"
		/tissue_type="pancreas"
		/clone_lib="RIKEN full-length enriched mouse cDNA library"
		/dev_stage="10 day old"
	misc_feature	5..1251
		/note="glycine C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase) (MGI:1349389) putative"
ORIGIN		
Query Match	5.6%; Score 175.2; DB 4; Length 1677;	
Best Local Similarity	49.9%; Pred. No. 8.5e-35;	
Matches	475; Conservative 0; Mismatches 468; Indels 9; Gaps 1;	
Qy	158 CAACAATTATTTAGGGCTCCAGCGATAGACGTTTGATCGATGCGCCCAACAGCAT	217
Db		
206 CCAATAACTACCTGGCGCTGAGCAGCCACCTCGAGTGATCCAGGCTGCGAGACTC	265	
Qy	218 TGCAGCAATTGGGACAGGAGCGGTTCACGTTTAAAGCAGAGCAATTCGCTCGGC	277
Db		
266 TGGAGGAGTTTGGAGCTGGACTCAGTTCTACTCGATTTATCTGTGGGACCCAGAGCATCC	325	
Qy	278 ATGAAAAGCTAGAAAAGAAATGCCAGCTTTAAACTGACAGAGCGCCCTGCTGTTTT	337
Db		
326 ATAGATCTAGAGCCAGATAGTACCCACCTTCACAGCGTGAGAGCCCATCTCTATC	385	
Qy	338 CGACGGGTACTTGGCCAAATGTCGGTGTCTTTCATCTCTGCCAGAAAGGAAGATGCA	397
Db		
386 CCAGCTGTTTCGATGCCAACCGCTGGCTATTTGAGGCTTTGCTGACCCAGAGATCGG	445	
Qy	398 TTTTAAGTGACCACTCAATCATGATGATGCGAGGCTGCGGACTTCTTAAGGCTG	457
Db		
446 TTCTGTGAGATGAGTGAACCATCTCTCCATCATTTGACGGCATCCGACTGTGCAAGCCC	505	
Qy	458 ATACAGTTGTTTATCGGCATATTGATGATGATCTTGAAAACAAAGCTGAATGAACAC	517
506 ACAAGTACCGCTTACCGCCACCTGGACATGCTGACCTGGAAGCCAGCTGAAGAGGCTC	565	
Qy	518 AGCGTTATCAGCGCGCTTTTATTCGTAAACAGACGAGGATTTTCAGCATGGATGGCAATCG	577
Db		
566 AGAAGCACAGGCTCGCGCTGGTGGCCACAGATGAGAGCTTTCTCCATGGATGGTACATCG	625	
Qy	578 CCCTCTTTGATCAGATCATCTCACTTTCGGAAACGCTCATGCTTCGTTGCTGCTTGATG	637
Db		
626 CTCCCTTACAGGACATCTCGCGCTTCGCCCCCAATATGGTGCCTGGCTTTTGTGGATG	685	
Qy	638 ATGCCACGCAACAGGAGTTTTCGGCGATTCGGGACAAAGAACGAGTGAATACATTTGGTG	697
Db		
686 AATGCCATGCCATGGCTTTCTTGGGCCACAGACGCGGGCACAGATGAGCTGTAGGTG	745	
Qy	698 TTTGTCCCGACATTTTATTCGGCACCTTAAGCAAG-----CTGTGGCGCGGGAAG	748
Db		
746 TCATGGACACAGGTCAACCATCATCAACTCCACCCTGGGGAAGGCGCTGGGTGGAGCATCAG	805	
Qy	749 GAGGTTTTCGGCAGGATCAGCGGTCTTCATCGACTTTTTCCTGAACCATGCCAGAACAT	808
Db		
806 GGGGATATACACCGGCGCTGAGCGCACTGGTGTCCCTGCTACGGCAGCGTTCTCGGCCCT	865	
Qy	809 TTATCTTTCAAAACCGCTATTTCGCCAGCCAGCTGTGCGGCTGCTCAGAGGCTTTCAACA	868
Db		
866 ACCTCTTCTCAACAGCGCTGCCCTGCTGTAGTCGGCTGTCTTAAGGCCCTGGACC	925	
Qy	869 TCATTGAAGCCAGCAGGGAAGAACAGCAGCTTTTATTTTATATCAGCATGATCAGAA	928
Db		
926 TGCTCATGAGAGCAATGCCATCATCAGTCTATGCTGTCGCAAGACCGCGGGTTCGCA	985	
Qy	929 CCAGTCTGAAGATATGGGTTATGCTGTGAAGAGAGATCACACCCGATTTATTCCTGTAG	988
Db		
986 GTAAGATGGAAGCTGCTGGTCTTACTGCTCAGAGAGCGGATCACCCCATCTGCGCCCTGGA	1045	
Qy	989 TCATTGGCGATGCCCATATAAACCGTCTATTGCTGAAAAAATCTCAGGCGCAAGGGAATTT	1048
Db		
1046 TGCTGGAGATGCAGGTTGCTCTTCAATGGCAGATGACATGTTAAAGAAAGTATCT	1105	
Qy	1049 ATGCTCTCGCATTCGCGCCGCCAACCGTTGCGCGGTGAAAGCCGATTCG 1100	
Db		
1106 TTGTCATCGGATTTAGTACCTCTGTGCTCCCTAAGGCAAGGCGCCGATCCG 1157		
RESULT 4	AY421131	1064 bp DNA linear GSS 17-DEC-2003
LOCUS	Homo sapiens GCAT gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.	
DEFINITION	AY421131	
ACCESSION	AY421131.1 GI:39777088	
VERSION	GSS.	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 1064)	
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.	
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios	
JOURNAL	Science 302 (5652), 1960-1963 (2003)	
PUBMED	14671302	
REFERENCE	2 (bases 1 to 1064)	
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,	

Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them
based on alignment.

FEATURES
source 1..1064
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1064
/gene="GCAT"
/locus_tag="HCM7455"

ORIGIN

Query Match 5.5%; Score 173.6; DB 10; Length 1064;
Best Local Similarity 49.8%; Pred. No. 2e-34;
Matches 474; Conservative 0; Mismatches 469; Indels 9; Gaps 1;

Qy 158 CAACAATATTTAGGCTCGCAAGCATAGAGTTTGATGATGCGACGCCCAACAGCAT 217
Db 19 CCAACAACTACCTGGGCTCGAGCAGCCACCCTGAGGTGATCCAGGCAGGTCTGCAGGCTC 78
Qy 218 TGCAGCAATTTGGACAGGAAGCAGCGGTTACGTTTAAACGACAGCAATTCGTCGTGGC 277
Db 79 TGGAGGAGTTTGGAGCTGGGCTCAGCTCTGTCCGCTTTATCTGTGGAACCCAGAGCATCC 138
Qy 278 ATCAAAAGCTAGAAAAGAGATTGCCAGCTTTAACTGCACAGAGCGGCCCTGCTGTTTT 337
Db 139 ACAAGATCTAGAAGCAAAATAGCCGCTTCCACAGCGGAGGATGCCATCTCTATC 198
Qy 338 CGAGCGGTTACTTGGCCAAATGTGGTGTCTTTTCATCTTCCAGCAAGAAAGAGATGCA 397
Db 199 CCAAGCTGTTATGACGCAAGCGCGGCTCTTTGAGGCGCTCTGACCCAGAGAGCAGCAG 258
Qy 398 TTTTAAGTACCAGCTCAATCAGTATGATCGAGCGCTGCGAGCTTTCTTAAGGCTG 457
Db 259 TCCTGTGGAGAGCTGAACCATGCTCCATCATCGACGGCATCCGGCTGTGCAAGGCC 318
Qy 458 ATACAGTTGTTTATCGGCATATTGATGAATGATCTTGAACCAAGCTGAATGAACAC 517
Db 319 ACAAGTACCGCTATGCGCACTGGACATGCGCGACCTAGAACCAAGCTCAGAGGCC 378
Qy 518 AGCGTTATCAGCGCGGTTTATCGTAACAGACGAGATTTACGATGGATGGACCAATCG 577
Db 379 AGAAGCATCGCTCGCGCTGTGSCCACTGATGGGCTTTTCCATGGATGGCGACATCG 438
Qy 578 CCCTCTTGATCAGATCATCTCACTTCGGAACGCTATCGCTTCGTTGGTGTTCATG 637
Db 439 CACCCCTGCAGGAGATCTGCTGCCTCTAGATATGGTGCCTTGTCTTCATGGATG 498
Qy 638 ATGCCACGCAACAGGAGTTTGGGCGATTTCGGGCAAGCAAGAGTGAATCTTTGGTG 697
Db 499 AATGCCATGCACTGGCTTCTGGGGCCACAGACGGGCAACAGATGAGTGTGGTG 558
Qy 698 TTTGTCCCGCATTTGTTATCGGCACCTTAAGCAAG-----CTGTTGGCGCGAAG 748
Db 559 TGAATGACCAAGTCAACATCATCAACTCCACCTTCGGGGAAGGCCCTGGGTGGAGCATCAG 618
Qy 749 GAGTTTGGCGGAGGATCAGCGTCTTCATCGACTTTTGTGTAACCATGCCAGAACAT 808
Db 619 GGGGCTACACGACAGGGGCTTGGGCGCTTGTGTCCTCTGCTGCGGACGCGCGCGGCAT 678
Qy 809 TTATCTTTTCAAAACCGCTATTCCGCGACCGAGCTGTGGGCTGTCTCAGAGGCTTTCAACA 868
Db 679 ACCTCTTCTCAAGACTGTGGCACCTGCTGTGGTGGCTTCCAGGCCCTAGATC 738
Qy 869 TCATTGAAGCCAGCAGGGAAGAAACGACAGCTTTTATTTTCTTATATCAGCATGATCAGAA 928
Db 739 TGCTGATGGGAGTAAACACCATTTGCTCAGTCTATGGCTGCCAAGACCCAGAGGTTCCGTA 798
Qy 929 CCAGTCTGAAGATATCGGTTATGTGGTGAAGAGATCACACCCGATTTATCTCTGTAG 988
Db 799 GTAAGATGGAAGCTGCTGGCTTCACTATCTCGGAGCCAGTCACCCCATCTGCGCTGTGA 858

Qy 989 TCATTGGCGATGCCCAATAAACGGTCTCTATTGCTGAAAAAAGTGCAGGCAAGGAATTT 1048
Db 859 TGGTGGTGGATGCGCGGCTCTCGCATGGCGGATGACATGCTGAAGAGGCACTCT 918
Qy 1049 ATGCTCTGCCATTGGCGGCCCAACGGTTGCGCGCGGTGAAGCCGGAATTCG 1100
Db 919 TTGTCATCGGGTTTCAGCTACCCCGTGTGCCCAAGGCAAGGCCCGGATCCG 970

RESULT 5

AK013138

LOCUS

DEFINITION

AK013138 2327 bp mRNA linear HTC 03-APR-2004
Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:281042022 product:glycine
C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase),
full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK013138
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

JOURNAL

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216]

COMMENT Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATTCGAGTTAATTAATATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SacI. Cloning sites, 5' end: XhoI; 3' end: SacI. Host: SclI.

FEATURES

source

Location/Qualifiers

1. .2327
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2810422022"
/db_xref="taxon:10090"
/clone="2810422022"
/tissue type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10, 11 days embryo"
16. .1266
/note="unnamed protein product; glycine C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase) (MGD|MG1:1349389)
putative"

CDS

/codon_start=1
/protein_id="BAB28673.1"
/db_xref="GI:12850315"
/translation="MWSAFMWHGALSFGRRASALQRLCILDSELEGIRGAGTWKSE RVITSROGSPRVGDSIGILNFCANNVGLSHSPAVIOAGLOTLEEFAGLSSTRI CGTOSLHNLEAKTAHOREDAIYPSCFDANAGLFEALITPEDAVLSDELNASHII DGILCKAHKRYRHLDMADEAKLEAKRHLRVATDGFSDMGDIAFLQDLCRLA AQYCALVFECHYATGFLGPTGRGTDELLGVMDQVITINSTLGNALGGAGGHTTGE PLVSLLRQSRPFLFSNLSPPAVVGCASKALDLMESNAIIQSMAAKTRFRSRKMEAA GFTVSGADHPICPVMLGDARLSSQWADMLKKGIFVIGFSYPVVPVPGKARIRVQISAV HSEBDIDRCVEAFVEVGRHLGALP"

ORIGIN

Query Match 5.5%; Score 173.6; DB 4; Length 2327;
Best Local Similarity 49.8%; Pred. NO. 2.5e-34;
Matches 474; Conservative 0; Mismatches 469; Indels 9; Gaps 1;
QY 158 CARACAAATTATTAGGGTCGCAGGATAGACGTTTGATCGATCGACCCAAACAGCAT 217
DB 221 CCAATAACTACTCTGGGCTGAGCAGCCACCTGCGAGTCCAGCGAGTCTGAGACTC 280
QY 218 TGCAGCAATTGGGACAGGACGCGGTTCACGTTTAAACGACAGGCAATTCGGTCTGGC 277
DB 281 TGGAGGAGTTTGGAGCTGAGCTCAGTTCTACTCGATTATCTGTGGGACCCAGAGCATCC 340
QY 278 ATGAAAAGCTAGAAAAGAAATTGCCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTT 337
DB 341 ATAAGAATCTAGAACCCAGATAGGCCACTTCCACGAGCGTGGAGCGCCATCCTCTATC 400
QY 338 CGAGCGGTACTTGGCCCAATTCGGTGTCTTTCATCTCCGAGAAAGAAAGATGCA 397
DB 401 CCAGCTGTTTCGATGCCCAACGCTGCGCTTATTGAGGCTTTCTGACCCAGGAGATCGG 460
QY 398 TTTTAAAGTGACAGCTCAATCATGCAAGTATGATCGACGGCTGCGGACTTCTTAAAGGTG 457

DB 461 TTCTGTGATGAGCTGAACCATGCTTCCATCTGACGGCATCCGACTGTGCAAGGCC 520
QY 458 ATACAGTTGTTTTCGCGATATGATATGATGATCTTGAACAAGCTGATGAACAC 517
DB 521 ACAAGTACCGTTACCGCCACCTGGACATGGCTGACCTGGGAAGCCAAAGCTGAAGGAGGCTC 580
QY 518 AGCGTTATCAGCGCGCTTTTATCCTAAACAGACGAGGATTTACAGCATGATGGCAACATCG 577
DB 581 AGAAGCACAGGCTCGGCTGGTGGCCACACAGATGAGGCTTTCTCCATGATGGTGCATCG 640
QY 578 CCCCTCTTGTATCAGATCATCTCACTTCGGAACGCTATCATGCTTCGTGCTGTTGATG 637
DB 641 CTCCTCTACAGGACATCTGCGCGCTTCGCGCCCAATATGGTGCCCTGCTTGTGGATG 700
QY 638 ATGCCACGACACAGGAGTTTGGGCGATTCGGGACAGGACGAGTGAATACTTGGTG 697
DB 701 AATGCCATGCCACTGGCTTTCTTGGGGCCACAGACGGGGCACAGATGAGTGTAGGTG 760
QY 698 TTTGTCCCGACATTTGTTATCGGCACCTTAAGCAAG-----CTGTTGGCGCGGAAG 748
DB 761 TCATGGACAGGTCACCATCATCACTCCACCTGGGGAGGCGCTGGGTGAGCATCAG 820
QY 749 GAGTTTTTGGCGCAGGATCAGCGTCTTCAATCGACTTTTTCGCTGAACCATGCCAGAACAT 808
DB 821 GGGGACATACACCGCGCTGAGCCACTGGTGTCCCTGCTACGGCAGCGTTCTCGGCCCT 880
QY 809 TTATCTTTTCAACCGCTATTTCGCGCAGCAGCTGTGCGGCTGCTCAGCAGGCTTTCAACA 868
DB 881 ACCTTTCTTCCACAGCGCTGCCCTGCTGTAGTCGCGCTGTGCTTCTTAAGGCCCTTGGACC 940
QY 869 TCATTGAAGCAGCAGGAGGAAAAACGACAGCTTTTATTTTCTTATATCAGCATGATCAGAA 928
DB 941 TGCTCATGGAGAGCAATGCCATCATCAGTCTATGGCTGCCAAGACCCGGCGTTCCGCA 1000
QY 929 CCAGTCTGAAAGATATGCGGTTATGTGTGAAAGAGATACACACCGATTATTCCTGAG 988
DB 1001 GTAAGATGGAAGCTGCTGTTTACTGTCTCAGGAGCGGATCACCCCATCTGCGCTGTGA 1060
QY 989 TCATTGGCGATGCCATAAAACCGTCTATTTGCTGAAAACCTCAGGGCAGGCAAGGATTT 1048
DB 1061 TGCTGGGAGATGCAACGTTGCTTCTCAATGGCAGATGACATGTTAAAGAAAGGATATCT 1120
QY 1049 ATGCTCTGCAATTCGGCCCGCAACCGCTGCGCGGTTGAAAAGCCGATTCG 1100
DB 1121 TTGTCATCGATTAGTACCTGTGTCTCCTAAGGSCAAGGCCCGGATCCG 1172

RESULT 6

AK076469 1737 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched library, clone:4833415L22 product:(2-amino-3-ketobutyrate-coenzyme A ligase), full insert sequence.
DEFINITION AK076469
ACCESSION AK076469
VERSION AK076469.1 GI:26096824
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
2
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

11042159
3
PUBMED
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitaura, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuura, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
4
JOURNAL
PUBMED
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
JOURNAL
PUBMED
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1737)
JOURNAL
PUBMED
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayaehida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saich, R., Saich, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp).
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>.
Location/Qualifiers
1. 1737
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:4833415L22"
/db_xref="taxon:10090"
/clone="4833415L22"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
35. 1499
misc_feature
/note="glycine C-acetyltransferase
(2-amino-3-ketobutyrate-coenzyme A ligase)
(MGDI:1349389)
putative"
ORIGIN
Query Match 5.4%; Score 171.4; DB 4; Length 1737;
Best Local Similarity 50.4%; Pred. No. 8.9e-34;

Matches 452; Conservative 0; Mismatches 436; Indels 9; Gaps 1;
QY 158 CAACAAATATTTAGGCTCGCAAGCATAGAGATTTGATCGATCGACCCCAACAGCAT 217
DB 239 CCATTAATACCTGGGCTGAGCAGCCACCTCCAGTGATCCAGGAGGTCTGAGACTC 298
QY 218 TGAGCAATTTGGGACAGGAGAGAGCGGTTCAGGTTTAAACAGACGGAATTCGTCGTC 277
DB 299 TGGAGGAGTTTGGAGCTGGGACTCAGTTTCTCATCTCGATTTATCTGTGGGACCCAGAGCATCC 358
QY 278 ATGAAAGCTAGAAAAGAGATTTGCCAGCTTTAAACTGACAGAGCGGCTGCTGTTTT 337
DB 359 ATAAGATCTAGAAAGCCAGATAGCCACCTTCCACGAGCTGAGAGGCCATCTCTATC 418
QY 338 CGAGCGTTACTTTGGCCAAATGTGGGTGTCCTTTTCATCTTCCAGAAAGAAAGATGTC 397
DB 419 CCAGCTGTTTCGATGTCACACGCTGGCCTATTTGAGGCTTTTGTGACCCAGAGATCGG 478
QY 398 TTTAAGTACACAGCTCAATCATGCAAGTATGATCGAGCGGTGCGGACTTTCTAAGCTG 457
DB 479 TTCTGTGAGATGAGCTGAACCATGCTTCCATTCATTTGACGGCATCCGACTGTGCAAGGCC 538
QY 458 ATACAGTTGTTTATCGGCATATTGATGATGAATGATCTTGAACCAACAGCTGAATCAAC 517
DB 539 ACAAGTACCGTTTACCGCCACCTGGAATGCTGACCTGGAAGCCAGCTGAAGAGGCTC 598
QY 518 AGCGTTATCAGCGCCGCTTTTATCGTAAACAGCGAGTATTTTCAGCATGGATGGCAATCG 577
DB 599 AGAAGCAGCGCTGGGCTGGTGGCCACAGATGAGGCTTTCTCCATGGATGGTGACATCG 658
QY 578 CCCTCTTTGATCAGATCATCTCACTTGGGAAAGCTATCATATGCTTCTGTTGTTGATG 637
DB 659 CTCCTCTTACAGGACATCTGCGGCTTGGCCGCCCAATATGTTGCTGCTGCTTTTGTGATG 718
QY 638 ATGCCACGCAACAGGAGTTTGGGCGATTTGGGACAGGAAACGAGTGAATCTTTGCTG 697
DB 719 ATGCGCATGCTGCTGCTTTCTTGGGCGCCACAGACCGGGGCAAGATGAATGCTGATG 778
QY 698 TTTGTCCCGACATTTGTTATCGGACCTTAAAGCAAG-----CTGTTGGCGCGGAAG 748
DB 779 TCATGACACAGGTACCATCATCACTCCACCTGGGAGAGGCGCTGGTGGAGCATCAG 838
QY 749 GAGTTTGGGAGGATCAGCGGTCTTCATTCGACTTTTGTGTAACCATGCCAGAACAT 808
DB 839 GGGGATATACACCGGCGCTGAGCCACTGGTGTCTCTGCTACGCGAGGTTCTCGGCCCT 898
QY 809 TTATCTTTCAACCGCTATTTCCGCGCAGCAGCTGTGGGCTGCTCAGAGGCTTTCAACA 868
DB 899 ACCTCTTCTCAACAGCCTGCCCCCTGCTGTAGTGGGCTGTGCTTCTAAGGCCCTTGHACC 958
QY 869 TCATTGAAGCGCAGCGGGAAGAAACGACAGCTTTTATTTTCTTATATCAGCATGATCAGAA 928
DB 959 TGCTCAGAGAGCAATGCCATCATCCAGTCTATGGCTGCCAGACCCGGGGTTCGCA 1018
QY 929 CCAGTCTGAAGAATATGGGTTATGTGTGAAGAGATCACAACCGATTATTCCTGTAG 988
DB 1019 GTAAGATGAAGCTGCTGTTTACTGTCTCAGAGCGCATCACCCCATCTGCCCCCTGTA 1078
QY 989 TCATTGGCGATGCCCAATAAAGCGTCTTATTTGCTGAAGAACTCAGGCGCAAGGAA 1045
DB 1079 TGCTGGAGATGACCGGTTGCTTTCTCAATGGCAGATGACATGTTTAAAGAAAGAA 1135

RESULT 7
LOCUS CR860400 1528 bp mRNA linear HTC 12-NOV-2004
DEFINITION Pongo pygmaeus mRNA; cDNA DKFZp459D1523 (from clone DKFZp459D1523).
ACCESSION CR860400
VERSION CR860400.1 GI:55731639
KEYWORDS HTC.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS0AT005ZC07_T0435_1&c=7364.r.

```
FEATURES
  source
    1. .911
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DB006YD07"
      /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
      /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
  Query Match      5.1%; Score 161.4; DB 5; Length 911;
  Best Local Similarity 50.5%; Pred. No. 3.4e-31;
  Matches 421; Conservative 0; Mismatches 404; Indels 9; Gaps 1;

Qy 150 CTGCTCTCAAAACAATTATTTAGGGCTCGCAAGCGATAGAGCTTTGATCGATGAGGCCA 209
    |||||
Db 858 CTTTGTGCAACANCTACCTGGCCCTGACAGCCACCCTGAGGTGATCCAGGCGAGTCT 799
    |||||

Qy 210 AACAGCATTTGAGCAATTTGGGACGAGGAGCGGTTACGTTTAAAGCAGGCAATTC 269
    |||||
Db 798 GCAGGCTCTGGAGAAGTTTGGAGCTGGCCCTCAGCTCTGTCCGCTTTATCTGTGAACCCA 739
    |||||

Qy 270 GGTCTGCNTGAAAGCTAGAAAGAGATTTGCCAGCTTTAACTGACAGAGCGGCCCT 329
    |||||
Db 738 GAGCATCCACAAGAACTTAGAAGCAAAATAGCCGCTTCCACCAGCGGGAGGATGCCAT 679
    |||||

Qy 330 GCTGTTTTCGAGCGGTACTTTGGCCAATGTGCGGTGCTCTTTCATCTTGCAGAAAGGA 389
    |||||
Db 678 CCTCTATCCAGCTGTTATGAGGCCAAGCCGGCTCTTTGAGGCCCTCTGTGACCCAGA 619
    |||||

Qy 390 AGATGTCAATTTAAGTGACCAAGCTCAATCATATGCAAGTATGATGACGGCTGCCAGTTTC 449
    |||||
Db 618 GGAGCGAGTCTGTGCGAGGAGCTGAACCATGCTCCATCATGACGCGCATCCGGCTGTG 559
    |||||

Qy 450 TAGGCTGTATACATTTGTTATCGGCATTTATGATATGAATGATCTTGAACAAAGCTGAA 509
    |||||
Db 558 CAAGGCCCAACAAGTACCCTATATGCCACCTTGGACATGCGCGACCTAGAACCAAGCTGCA 499
    |||||

Qy 510 TGAACAACAGCGTTATCAGCGCGCTTTTATCGTAAACAGACGAGTATTTACAGATGATGG 569
    |||||
Db 498 GGAGGCCCAAGAGATCGCTGCGCCCTGTGTCGACATGATGGGGCCCTTTCCATGATGG 439
    |||||

Qy 570 CACAATCGCCCTCTTTGATCAGATCATCTCACTTTCGAAACGCTATCATGCTTCGTGGT 629
    |||||
Db 438 CGACATCGCACCCCTGCGAGGATCTGCTGCTCGCCTCTAGATATGTTGCTGCTGCTT 379
    |||||

Qy 630 CGTTGATGATGCCCAAGCAAGAGTTTGGGCGATTCGGGACAGGAACGAGTGAATA 689
    |||||
Db 378 CATGATGATATGCCATGCCACTGTGCTTCTGGGGGCCCAAGGACGGGGCCAGATGAGCT 319
    |||||

Qy 690 CTTTGTGTTTGTCCGACATTTGTTATCGGCACCTTAAGCAAGCT-----GTTGG 740
    |||||
Db 318 GCTGGTGTGATGACCAAGTACCATCATCATCACTCCACCTTGGGAGGCCCTTGGTGG 259
    |||||

Qy 741 CGCGGAAGGAGGTTTTCGGCAGGATCAGCGGTCTTCATCGACTTTTTCGTGAACCAATGC 800
    |||||
Db 258 AGCATCAGGGGGCTACACGACAGGCGCTGGGCCCTGTGTCTCTGTGCGGACGCGC 199
    |||||

Qy 801 CAGAACATTTATCTTTAAACCGCTATTCGCCAGCAGCTGTGCGGGTGTCTCAGAGGC 860
    |||||
Db 198 CCGGCCATACCTCTCTCCAAACAGTCTGCCACCTGCTGTGCTGGCTCGCCTCCAAGGC 139
    |||||

Qy 861 TTTCAACATCATTTGAAGCCAGCGGAACAGACAGCTTTTATTTCTTATATCAGAT 920
    |||||
Db 138 CCTAGATCTGCTGATGGGGAGTAAACACCATTTGCTCAGTCTATGCTGCTGCAAGACCCAGAG 79
    |||||

Qy 921 GATCAGAACCAGTCTGAAGAATATATGGGTTATGTGTTGAAAGGAGATCACACACC 974
    |||||
```

```
Db 78 GTTCGCTAGTAAGATGAAGCTGCTGCTTCACTTTTTCGAGGCCAGTCACC 25
    |||||
RESULT 9
CA778095 794 bp mRNA linear EST 03-DEC-2002
LOCUS ip17h09.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6217601 3'
DEFINITION Similar to SW:KBL HUMAN 075600 2-AMINO-3-KETOBUVRATE COENZYME A
LIGASE, MITOCHONDRIAL PRECURSOR ;, mRNA sequence.
ACCESSION CA778095
VERSION CA778095.1 GI:26015970
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE 1 (bases 1 to 794)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
  Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
  Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A.,
  Schmitt,A., Theising,B., Ritter,S., Ronko,I., Bennett,J.,
  Cardenas,M., Gibbons,M., McCann,K., Cole,R., Tsagareishvili,R.,
  Williams,T., Jackson,Y. and Bowers,Y.
  Endocrine Pancreas Consortium
  Unpublished (2000)
  Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
  Endocrine Pancreas Consortium
  Harvard University, Howard Hughes Medical Institute
  Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
  MA 02138
  Tel: 617-495-1812
  Fax: 617-495-8557
  Email: dmelton@biohp.harvard.edu
  Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
  Washington University Genome Sequencing Center For information on
  obtaining a clone please contact: Dr. Hiroshi Inoue
  (hinoue@im.wustl.edu)
  Possible reversed clone: similarity on wrong strand
  Seq primer: -40UP from Gibco
  High quality sequence stop: 439.
FEATURES
  source
    1. .794
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:6217601"
      /tissue_type="Purified pancreatic islet"
      /lab_host="DH10B"
      /clone_lib="HR85 islet"
      /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
      NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
      Size-selected on agarose gel. Average insert size ~1kb. 5'
      XhoI site was destroyed after directional cloning.
      Amplified once. Contact information: Hiroshi Inoue, MD,
      Metabolism Div. (Alan Permutt Lab), Washington University
      School of Medicine, Box 8127, 660 South Euclid Ave., St.
      Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
      314-362-1916, Fax: 314-747-2692."
ORIGIN
  Query Match      5.1%; Score 159.8; DB 6; Length 794;
  Best Local Similarity 54.1%; Pred. No. 8.8e-31;
  Matches 326; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

Qy 158 CAAACCAATTATTTAGGGCTCGCAAGCGATAGAGCTTTGATCGATGCCCAACAGCAT 217
    |||||
Db 124 CCACCACTACTCTGGGCTCGAGCAGCCACCTGAGGTGATCCAGGCTCTGAGGCTC 183
    |||||

Qy 218 TGACGCAATTTGGGACAGGAGCAGCGGTTCACTTTAACGACAGGCAATTCGGTCTGGC 277
    |||||
Db 184 TGGAGGAGTTTGGAGCTGGGCTCAGCTCTGTCCGCTTTATCTGTGGAACCCAGAGCATCC 243
    |||||
```

Qy 278 ATGAAAAGCTAGAAAAGAGATTGCCAGCTTTTAAACTGACAGAAGCGGCCCTGCTGTTTT 337
 Db 244 ACAGAGATCTAGAGCAAAATAGCCCGCTTCCACGAGGGAGATGCCATCTCTATC 303
 Qy 338 CGAGCGGTACTTGGCCCAATGCGGTGCTTTTATCCTTCCAGAAAAGGAAGATGTCA 397
 Db 304 CCAGCTGTTATGACGCCAAACCGGSCCTCTTTGAGGCCCTGCTGACCCAGAGACGCGAG 363
 Qy 398 TTTTAAGTGACAGCTCAATCATGCAAGTATGATCGAGCGGTGCGGACTTTCTTAAGGCTG 457
 Db 364 TCCTGTGCGGACGAGCTGAACCATGCTCCATCATGACGGCATCGGCTGTGCAAGGCC 423
 Qy 458 ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTGAATAAAGCTGAATGAACAC 517
 Db 424 ACAAGTACCGGTATGCGCCCTGACATGCGGACCTTAGAGCCAGCTGAGAGGCC 483
 Qy 518 AGCGTTATCAGCGCGGTTTATTCGTAACAGACGAGGATATTCAGCATGATGCGCAATCG 577
 Db 484 AGAAGCATCGGCTGCGCCTGCTGCGCACTGATGGGCCCTTTTCCATGATGCGGACATCG 543
 Qy 578 CCCCTCTTGATCAGATCATCTCACTTGGAAACGCTATCATGCTTGGTGTGCTGATG 637
 Db 544 CACCCCTGCGAGGATCTGCTGCTCGCTCTAGATATGCTGCTGCTTCTATGATG 603
 Qy 638 ATGCCCGCAACAGGAGTTTGGCGGATTTCGGACAAAGAACGAGTGAATCTTTGGTG 697
 Db 604 AATGCCATGCACTGCGCTCTCTGGGGCCACAGACGGGGACAGATGAGCTGCTGGTG 663
 Qy 698 TTTGTCCCGACATGTTATTCGCGCACCTTAAAGCAAGCTGTTGGCGCGGAAGAGGTTTG 757
 Db 664 TGATGGACCGGTCAACCATCATCAACTCCACCCTGGGGAGGCCCGGGTGGAGCATCAG 723
 Qy 758 CGG 760
 Db 724 GGG 726

RESULT 10

CA778022
 LOCUS
 DEFINITION
 ip16h09.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6217600 3',
 similar to SW:KBL HUMAN O75600 2-AMINO-3-KETOBUVRATE COENZYME A
 LIASE, MITOCHONDRIAL PRECURSOR ;, mRNA sequence.

CA778022
 CA778022.1 GI:26015897

EST.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 810)

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Clifton, S.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
 Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarelis, V.,
 Williams, T., Jackson, Y. and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)

Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 439.

FEATURES

source

1. 810
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6217600"
 /cissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /notes="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:
 NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 5.1%; Score 159.8; DB 6; Length 810;
 Best Local Similarity 53.2%; Pred. No. 8.9e-31;
 Matches 366; Conservative 0; Mismatches 313; Indels 9; Gaps 1;

Qy 158 CAACAATATTATTAGGCTCGCAAGCGATAGAGTTTGTATCGATCGACCCCAACAGCAT 217
 Db 122 CCACAACTACTCTGGGCTGAGCAGCCACCTGTAGGTGATCCAGGCGGTCTGAGGCTC 181
 Qy 218 TGCAGCAATTTGGGACAGGAGCGGTTTACGTTTAAACGACAGGCAATTCGGTCTGGC 277
 Db 182 TGGAGGAGTTTGGAGCTGGGCTCAGCTCTGCTCGCTTTTATCTGTGGAAACCCAGAGCATCC 241
 Qy 278 ATGAAAAGCTAGAAAAGAGATTGCCAGCTTTAAACTGACAGAACCGGCCCTGCTGTTTT 337
 Db 242 ACAAGATCTAGAAGCAAAATAGCCCGCTTCCACAGCGGAGGATGCCATCTCTATC 301
 Qy 338 CGAGCGGTACTTGGCCCAATGTCGGTGTCTTTTCATCTTGGCCAGAAAAGGAAGATGCA 397
 Db 302 CCAGCTGTTATGACGCCAACCGCGGCTCTTTGAGGCCCTGCTGACCCAGAGACGCGAG 361
 Qy 398 TTTTAAGTGACACGAGCTCAATCATCATGATGATGACGAGCGGTGCGGACTTCTTAAGGCTG 457
 Db 362 TCCTGTCGAGAGAGCTGAACCATGCTCCATCATCGACGGCATCCGGCTGTGCAAGGCC 421
 Qy 458 ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTGAAAAACAAGCTGAATGAACAC 517
 Db 422 ACAAGTACCGCTATCGCCACCTGGACATGGCGGACCTAGAAGCCAAGCTGCAGAGGCC 481
 Qy 518 AGCGTTATCAGCGCGGTTTTTATCGTAACAGACGAGATTTTCAGCATGATGGCACAATCG 577
 Db 482 AGAAGCATCGGCTGCGGCTGGTGGCCACTGATGGGGCTTTTTCATGATGGCACAATCG 541
 Qy 578 CCCCTCTTGATCAGATCATCTCACTTCGGAACCGCTATCATGCTCTGCTGCTGTTGATG 637
 Db 542 CACCCCTGCAAGAGATCTGTCCTCGCTCTAGATATGTTGCTGCTGCTGCTGCTGATG 601
 Qy 638 ATGCCCAACGCAAGGAGTTTTTGGGGGATTCGGGACAAAGAACGAGTGAATCTTTGGTG 697
 Db 602 AATGCCATGCCATGCGCTTCTTGGGGGCCACAGGACGGGGCACAGATGAGCTGCTGGGTG 661
 Qy 698 TTTGTCCCGACATGTTTATCGGCACCTTAAGCAAG-----CTGTTGGCGGGAAG 748
 Db 662 TGATGGACCAAGGTCAACCATCATCAACTCCACCTTGGGGAAAGGCCCTGGGTGGAGCATCAG 721
 Qy 749 GAGGTTTTGCGGAGGATCAGCGGTCTTCATCGACTTTTTTGTCTGAACCATGCCAGAACAT 808
 Db 722 GGGGCTACAGACAGGGGCTTGGGCCCTGGTGTGCTGCTGCGGACGGCGCCCGGCAT 781
 Qy 809 TTATCTTTCAAAACCGCTATTTCGCCAGC 836


```
Db 19 CCAACAACACTACCTGGGCTGAGCAGCCACCCTGAGGTGATCCAGGCGAGGTCTGCAGGCTC 78
Qy 218 TGCAGCAATTTGGGACAGGAGCGGTTTACGCTTTAAACGACAGGCAATTCGGTCTGGC 277
Db 79 TGGAGCAGTTTGGAGCTGGCCCTAGCTCTGTCCGCTTTATCTGTGGAAACCCAGAGCATCC 138
Qy 278 ATGAAAAGCTAGAAAAGAAGATTGCCAGCTTTAAACTGACAGAAAGCGGCCCTGCTGTTTT 337
Db 139 ACAAGAACTTAGAAGCAAAATAGCCGCTTCCACAGCGGAGGATGCCATCCTCTATC 198
Qy 338 CGAGCGTTACTTGGCCAAATGTGGGTCTCTTTTCATCTCTTGCAGAAAAGGAAGATGCA 397
Db 199 CCAGCTGTTTGGAGCAACCGCGGCTCTTTGAGGCCCTGCTGACCCCGGAGACGAC 258
Qy 398 TTTTAAGTGACAGCTCAATCATCAAGATGATGCGAGCGCTGCGACTTTCTTAAGGCTG 457
Db 259 TCCGTGCGAGAGCTGACCATGCCCTCCATCATCGAGCGGATCCGGCTGTGCAAGGCC 318
Qy 458 ATACAGTTGTTTATCGGCATATTGATATGATGATCTTGAACCAAGCTGAATGAACAC 517
Db 319 ACAAGTACCGCTATCGCCACCTGACATGGCGACCTNNNGCCAAAGCTGCNNNNNNN 378
Qy 518 AGCTTATCAGCGCGGTTTATCGTAACAGAGGAGATTTCAGCATGATGGCAATCG 577
Db 379 NNAAGCATCGGCTGCGGCTGGTGGCCACTGATGGGGCTTTTCCATGGATGGCGACATCG 438
Qy 578 CCCCTCTTGTATCATCATCTCACTGCGAAACGCTATCATGCTCTGCTGCTGCTGATG 637
Db 439 CACCCCTGCGAGGATCTGCGGCTCGCTTAGATATGGCGGCTGCTGCTGCTGATG 498
Qy 638 ATGCCACGCAACAGGAGTTTGGGCGATTGGGACAGGAAACGAGTGAATCTTTGGTG 697
Db 499 AATGCCATGCCACTGGCTTCTGGGGCCACAGGACGNNNNNNNNNNAGCTGCTGGTG 558
Qy 698 TTTGTCGCCGATTTGTTATCGGCACTTAAGCAAG-----CTGTTGGCGCGGAAG 748
Db 559 TGATGGACAGGTCACCATCATCAACTCCACCTGGGGGAAGGCCCTGGTGGAGCATCAG 618
Qy 749 GAGGTTTGGCGCAGGATCAGCGGTCTTCATCGACTTTTGTGAACCATGCCAGACAT 808
Db 619 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 678
Qy 809 TTATCTTTCAAAACCGCTATTCCGCGCAGCAGCTGTGGGCTGTCTCAGAGGCTTTCAACA 868
Db 679 ACCTCTTCCAAAGTCTGCCACTGTGTGTGGTGGCGGCTTCAAGGCCCTAGATC 738
Qy 869 TCATTGAAGCCAGCAGGGAAGAAACGACAGCTTTTATTTCTTATATCAGCATGATCAGAA 928
Db 739 TGCTGATGGGAGTAAACACCATNGTCCAGTCTATGGCTGCAAGACCCAGAGGTTCCGTA 798
Qy 929 CCAGTCTGAAGAAATATGGGTTATGTGGTGAAGGAGATCACACCCGATTTATCTCTAG 988
Db 799 GTAAGATGGAAGCTGTGGGCTTCACTATCTGGGAGCAGTCAACCCATCTGCCCCTGTA 858
Qy 989 TCATTGGCGATGCCCAATAAACCGTCTCTATTGCTGAAAACTGACGCGCAAGGGAATTT 1048
Db 859 TGCTGGGTGATGCCCGGCTGCGCTCTCGCATGGCGGATGACATGCTGAAGAGGCACT 918
Qy 1049 ATGCTCTGCAATTCGGCGCGCAACACCGTTGGCGCGGGTGAAGCCGATTCG 1100
Db 919 TTGTCATCGGGTTTCACTACCCGCTGGTCTCCCAAGGCGCAAGGCCCGATCCG 970
```

RESULT 14

BG752866

LOCUS

DEFINITION BG752866 787 bp mRNA linear EST 15-MAY-2001
602732509F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875915 5',
mRNA sequence.

ACCESSION

BG752866

VERSION

BG752866.1 GI:14063519

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE

1 (bases 1 to 787)

NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-rc@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LUCM1758 row: a column: 04

High quality sequence stop: 785.

Location/Qualifiers

1. .787

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4875915"

/tissue_type="normal pigmented retinal epithelium"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_43"

/note="Organ: eye; Vector: pOTB7; Site:1: XhoI; Site:2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library. |"

ORIGIN

Query Match 5.0%; Score 157.2; DB 2; Length 787;
Best Local Similarity 54.0%; Pred. No. 4.3e-10;
Matches 321; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 158 CAACAATATTATTAGGCTCGCAAGCGATAGAGTTTGATCGATGCAAGCCCAACAGCAT 217
Db 69 CCAACAATACCTTGGGCTCGAGCAGCCACCTCGAGGTGATCCAGGCGAGGTCTGCAGGCTC 128
Qy 218 TGCAGCAATTTGGGACAGGAGCGGTTTACGCTTTAAACGACAGGCAATTCGGTCTGGC 277
Db 129 TGGAGGATTTTGGAGCTGGGCTCAGCTCTGTCGCTTTTATCTGTGGAACCCAGAGCATCC 188
Qy 278 ATCAAAAGCTAGAAAAGAAAGATTGCCAGCTTTTAAACTGACAGAGCGGCCCTGCTGTTTT 337
Db 189 ACAAGATCTTAGAAGCAAAATAGCCGCTTCCACGACGGGAGGATGCCATCTCTATC 248
Qy 338 CGAGCGTTACTTGGGCAATGTGCGGTCTCTTTTCAATCTTGCAGAAAAGGAAGATGCA 397
Db 249 CCAGCTGTTATGACGCCAACGCCGCTCTTTGAGGCCCTGCTGACCCCAAGAGACGACG 308
Qy 398 TTTTAAGTGACAGCTCAATCATGCAAGATGATGAGCGGCTGCCGACTTCTTAAGGCTG 457
Db 309 TCCGTGCGGACGAGCTGAACCATGCTCCATCATCGACGGCATCCGGCTGTGCAAGGCC 368
Qy 458 ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTGAACCAAGCTGAATGAACAC 517
Db 369 ACAAGTACCGCTATGCGCACTGACATGCGCCACCTTAGAAGCCACAGCTGCAGAGGCC 428
Qy 518 AGCGTTATCAGCGCGGTTTATCGTAACAGAGGAGATTTCAGCATGATGGCAATCG 577
Db 429 AGAAGCATCGGCTGCGGCTGGTGGCCACTGATGGGCTTTTCCATGGATGGCGACATCG 488
Qy 578 CCCCTCTTGTATCATCATCTCACTGCGAAACGCTATCATGCTCTGCTGCTGCTGATG 637
Db 489 CACCCCTGCGAGGATCTGCTGCTCGCCTCTAGATATGGTGGCCCTGCTTCTTATGGATG 548

Qy 638 ATGCCACGCAACAGGAGTTTGGCGGATTTCGGGACAGGACGAGTGAATCTTGGTG 697
Db 549 AATGCCATGCGACCTGGCTTCTCTGGGGCCACAGGACGGGGCACAGATGAGTGTGGGTG 608
Qy 698 TTTGTCCCGACATGTTTATCGGCACCTTAAAGCAAGCTGTTGGGCGCGAAGGAG 751
Db 609 TGATGGACCAAGTCAACATCATCACTCCACCTGGGGAGGCCCTGGGTGGAG 662

RESULT 15
LOCUS DNO19683
DEFINITION JGI CHAR2120.fwd NIH XGC tropliv1 Xenopus tropicalis cDNA clone
IMAGE:7734748 5', mRNA sequence.

ACCESSION DNO19683
VERSION DNO19683.1 GI:59863317
KEYWORDS EST.

SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 871)
Richardson,P., Lucas,S., Rokheer,D., Dettler,J.C., Ng,D.C.,
Brokstein,P. and Lindquist,E.A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
Other ESTs: JGI CHAR2120.rev
Contact: Lindquist,E.A., Richardson,P.

DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710

Email: cdna@jgi-psf.org
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Bruce Blumberg Laboratory, University of
California, Irvine
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov

Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: CHAR 0021 row: o column: 2
High quality sequence stop: 863.

FEATURES
Location/Qualifiers
1..871
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7734748"
/sex="male"
/tissue_type="Liver"
/dev_stage="Adult"
/lab_host="Electromax DH10B T1 Phage Resistant cells"
/clone_lib="NIH XGC tropliv1"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The
library was prepared from 5 ug of poly A+ RNA by oligo-dT
priming
(5'-GAGAGAGAGAGAGAGAGACTAGTCTCGAGTTTTTTTTTTTTTTT-3')
and Stratascript reverse transcriptase. After ligation of
EcoRI adapters (5'-AATCGGACGAGG-3') followed by kinasing
adapters and by XhoI digestion, the cDNA was size selected
by chromatography on Sepharose CL-2B columns and fractions
containing cDNAs larger than 1000 bp were ligated into
EcoRI/XhoI-digested pCS107. Reference for library
construction: Current Genomics 4, 635-644. Library
constructed by Michelle Tabb and Bruce Blumberg (Dept of
Developmental and Cell Biology, University of California,
Irvine)."

ORIGIN

Query Match 5.0%; Score 157.2; DB 8; Length 871;
Best Local Similarity 53.9%; Pred. No. 4.5e-30;
Matches 371; Conservative 0; Mismatches 308; Indels 9; Gaps 2;

Qy 158 CAAACAAATTTATAGGCTCGCAAGCGATAGAGCTTTGATCGATCGAGCCCAACAGCAT 217
Db 177 CCAATAACTACTTAGGCTTTCCAGCCATCCCAAGGTGATCCGTCGCTGCTGAGACAC 236
Qy 218 TGCAGCAATTTGGGACAGGAGCGGTTTCACTTTAAACAGCAAGCAATTCGGTCTGGC 277
Db 237 TAGAGGAGTTTGGAGCGGCTCTCAGCTCAGTGAAGTTTATCTGTGGACACAGAATATCC 296
Qy 278 ATGAAAGCTAGAAAAGAAAGATTGCCAGCTTTAAACTGACAGAAAGCGCCCTGCTGTTT 337
Db 297 ATAAGAGTCTAGAAACAGAGAGATTGCCCGGTTCCACAGAGAGAGATGCCATCTTGTATA 356
Qy 338 CGAGCGGTTACTTTGGCCCAATGTGCGTGTCTCTTTCATCTTCCAGAAAAAGAGATGTCA 397
Db 357 TCAGCTCTTTGATGCCCAATGCTGTAATATTGAGGCTCTTCTGTCCCTGAAGATGCAG 416
Qy 398 TTTTAAGTGACCACTCAATCATGCAAGTATGATCGACGGCTGCCGACTTTTCTAAGGCTG 457
Db 417 TGCTCTCAGATGAGTTGAACACCGCTCAATCATTTGATGGTATCCGCTTTTGAAGGCCA 476
Qy 458 ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTTGAAAACAAGCTGAATGAACAC 517
Db 477 ATAAGTACCGCTATTAAGCACATGACCTAGATGATTTAGAAAGCCCAAGCTTTAAGGAGCC 536
Qy 518 AGCGTTATCAGCGCGCTTTTATGTTATGTTACAGACGAGGATTTTCAAGTATGATGGCAATCG 577
Db 537 AGAAGCACCGCTTGGCTCTCAITGGCCACAGCGGTGTTTTTCTATGACGAGACATTG 596
Qy 578 CCCTCTTGATCAGATCATCTCACTTGCAGAAAGCTATCATGCTTCTGTGGTCTGTTGATG 637
Db 597 CACCTCTTCGTGAGATTGCAACTTAGCAAGGCGCTACAATGCCCTTAGTCTTTATAGATG 656
Qy 638 ATGCCACGCAACAGGAGTTTGGGCGATTTCGGGACAGGAAAGAGTGAATACATTTGGTG 697
Db 657 AGTGCCATGCTACCGGTTTTTATGGACCCAAATGGAAGGGGACAGATGAGCTTCTCGGAG 716
Qy 698 T-----TTGTCCCGACATTGTTTATCGGCACCTTTAAGCAAAAGC---TGTTGGCGCGGAAG 748
Db 717 TCATGGATCAAGTGACAAATTTGTGAATTTACCTTTGCAAAAGCACTGGAGGGGCTGCAG 776
Qy 749 GAGGTTTTCGGGAGGATCAGCGGTCTTCAATCGACTTTTTTGTGTAACCATGCCCAGAACAT 808
Db 777 GTGGGTACACCACTGGGCCCAAGGCTTTGATTTGACTTACTACGGCAGCGCTCTCGGCCAT 836
Qy 809 TTATCTTTCAACCGCTATTTCGGCCAGC 836
Db 837 ACCTGTTTCTCAAAACACCTTTCCGCCAGC 864

Search completed: February 11, 2006, 23:19:32
Job time : 8022 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2006, 21:06:35 ; Search time 384 Seconds
(without alignments)
14609.335 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 3156

Sequence: 1 ggatccacgaggttacgagc.....atcggtatgatgcgaattc 3156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PPUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3156	100.0	3156	3	US-09-633-927-1
2	1612.6	51.1	8478	3	US-08-676-818-1
3	1612.6	51.1	8478	3	US-09-407-549-1
4	1612.6	51.1	8478	3	US-09-407-728-1
5	867	27.5	5793	3	US-09-869-855A-3
6	867	27.5	5943	3	US-09-869-855A-2
7	867	27.5	7336	3	US-08-418-085-3
8	867	27.5	7336	3	US-09-099-011A-3
9	867	27.5	7336	3	US-08-470-369-21
10	867	27.5	7336	3	US-09-098-877B-3
11	867	27.5	8119	2	US-08-460-343B-1
12	867	27.5	8119	2	US-08-398-028B-1
13	867	27.5	8119	2	US-08-504-265B-1
14	867	27.5	9144	3	US-08-556-978B-79
15	865.4	27.4	7325	9	5304637-21
16	853	27.0	5609	3	US-09-313-677-14
17	853	27.0	5609	3	US-09-313-677-15
18	853	27.0	7026	3	US-09-313-677-20
19	853	27.0	7344	3	US-09-313-677-16
20	762.4	24.2	1200	2	US-08-642-045B-3
21	762.4	24.2	1200	2	US-08-852-268-1
22	759	24.0	759	3	US-08-697-186B-10
23	703	22.3	6169	2	US-08-875-154-2
24	701	22.2	6837	3	US-09-928-847B-49

25	471	14.9	6405	3	US-09-869-865A-1	Sequence 1, Appli
26	336	10.6	5390	6	PCT-US96-12545-5	Sequence 5, Appli
c 27	286.6	9.1	3348	3	US-09-710-279-3541	Sequence 3541, Ap
c 28	286.6	9.1	3830	3	US-09-710-279-3934	Sequence 3934, Ap
c 29	286.6	9.1	3876	3	US-09-710-279-3895	Sequence 3895, Ap
30	285	9.0	984	3	US-09-134-001C-634	Sequence 634, App
31	285	9.0	999	3	US-09-710-279-539	Sequence 539, App
32	285	9.0	999	3	US-09-710-279-2561	Sequence 2561, Ap
33	284.4	9.0	1017	3	US-08-935-263-7	Sequence 7, Appli
34	284.4	9.0	1017	3	US-09-594-185-7	Sequence 7, Appli
35	284.4	9.0	1017	3	US-10-033-078-7	Sequence 7, Appli
36	284.4	9.0	1017	3	US-10-033-933-7	Sequence 7, Appli
37	255.2	8.1	1164	3	US-08-935-263-5	Sequence 5, Appli
38	255.2	8.1	1164	3	US-09-594-185-5	Sequence 5, Appli
39	255.2	8.1	1164	3	US-10-033-078-5	Sequence 5, Appli
40	255.2	8.1	1164	3	US-10-763-933-5	Sequence 5, Appli
c 41	249.6	7.9	3173	3	US-08-956-171B-338	Sequence 338, App
c 42	249.6	7.9	3173	3	US-08-781-986A-338	Sequence 338, App
43	184.8	5.9	1066	3	US-08-936-165A-110	Sequence 110, App
44	169.6	5.4	3792	2	US-08-992-334-1	Sequence 1, Appli
45	169.6	5.4	3792	3	US-08-302-752-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-633-927-1
; Sequence 1, Application US/09633927
; Patent No. 6656721
; GENERAL INFORMATION:
; APPLICANT: HOHMANN Dr., Hans-Peter
; APPLICANT: MOUNCEY Dr., Nigel J.
; APPLICANT: SCHUEKER Dr., Heinrich W.
; APPLICANT: STEBBINS Dr., Jeffrey W.
; TITLE OF INVENTION: PROCESS FOR PRODUCING A TARGET FERMENTATION PRODUCT
; FILE REFERENCE: Process for Prod. Target Ferm. Product
; CURRENT APPLICATION NUMBER: US/09/633,927
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3156
; TYPE: DNA
; ORGANISM: Bacillus subtilis
; US-09-633-927-1

Query Match	100.0%;	Score 3156;	DB 3;	Length 3156;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3156;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGTTAAACGAGCGGTTAGACAGA	60	
Db	1	GGATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGTTAAACGAGCGGTTAGACAGA	60	
Qy	61	ATGAAGAAGCCGCGGTACATCGTAACCTCGGTCATATGATGAGGAGCCCGGTTCCAGAG	120	
Db	61	ATGAAGAAGCCGCGGTACATCGTAACCTCGGTCATATGATGAGGAGCCCGGTTCCAGAG	120	
Qy	121	AGGAATTTGATGCGGAAATCAACGGTCTGTCCTCAACCAATTTATTTAGGCTCGCA	180	
Db	121	AGGAATTTGATGCGGAAATCAACGGTCTGTCCTCAACCAATTTATTTAGGCTCGCA	180	
Qy	181	AGCGATAGAGCTTTGATCGATGAGCCCAACAGCAATTTGGGACAGGAGGC	240	
Db	181	AGCGATAGAGCTTTGATCGATGAGCCCAACAGCAATTTGGGACAGGAGGC	240	
Qy	241	AGCGGTTTACGAGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATT	300	
Db	241	AGCGGTTTACGAGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATT	300	
Qy	301	GCACGCTTTAACTGACAGAGCGGCCCTGCTGTTTTCGAGCGGTTACTTGGCCCAATGTC	360	
Db	301	GCACGCTTTAACTGACAGAGCGGCCCTGCTGTTTTCGAGCGGTTACTTGGCCCAATGTC	360	

Db 301 GCCAGCTTTAAACTGACAGAAAGCGCCCTGCTGTTGTTTTTCGAGCGGTTACTTGGCCAAATGTC 360
Qy 361 GGTGTCCCTTTTCATCCTTTCGACGAAAGGAAGATGTCATTTTAAAGTGACCGAGCTCATATCAT 420
Db 361 GGTGTCCCTTTTCATCCTTTCGACGAAAGGAAGATGTCATTTTAAAGTGACCGAGCTCAATCAT 420
Qy 421 GCAAGTAGATGACGAGCTGCGGACTTCTTAAGGCTGATACAGTTGTTTATCGGCATATT 480
Db 421 GCAAGTAGATGACGAGCTGCGGACTTCTTAAGGCTGATACAGTTGTTTATCGGCATATT 480
Qy 481 GATATGAATGATCTTTGAAAAACAAGCTGAATGAAACAACAGCGTTATFACGCGCGCTTTTATC 540
Db 481 GATATGAATGATCTTTGAAAAACAAGCTGAATGAAACAACAGCGTTATFACGCGCGCTTTTATC 540
Qy 541 GTAACAGACGAGATATTACAGCATGATGCGACAAATCGCCCTCTTTGATCAGATCATCTCA 600
Db 541 GTAACAGACGAGATATTACAGCATGATGCGACAAATCGCCCTCTTTGATCAGATCATCTCA 600
Qy 601 CTTGCGAAACGCTATCATGCCCTTCTGTGCTCGTTGATGATGCCACGCAACAGGAGTTTTTG 660
Db 601 CTTGCGAAACGCTATCATGCCCTTCTGTGCTCGTTGATGATGCCACGCAACAGGAGTTTTTG 660
Qy 661 GCGGATTCGGGACAAAGGACGAGTGAATACCTTTGCTGTTGTCGCCGACATTTGTTATCGGC 720
Db 661 GCGGATTCGGGACAAAGGACGAGTGAATACCTTTGCTGTTGTCGCCGACATTTGTTATCGGC 720
Qy 721 ACCTTAAGCAAAAGCTGTTGGCGCGGAGGAGTTTTGCGGAGGATCATCAGCGGCTCTTCATC 780
Db 721 ACCTTAAGCAAAAGCTGTTGGCGCGGAGGAGTTTTGCGGAGGATCATCAGCGGCTCTTCATC 780
Qy 781 GACTTTTTGCTGAACCATGCCAGAACATTTATCTTTCAAACCGCTATTTCGCCAGCCAGC 840
Db 781 GACTTTTTGCTGAACCATGCCAGAACATTTATCTTTCAAACCGCTATTTCGCCAGCCAGC 840
Qy 841 TGTGCGGCTGTCTACGAGGCTTTCAACATCATTTGAAGCCAGCAGGAGGAAAAACGACAGCTT 900
Db 841 TGTGCGGCTGTCTACGAGGCTTTCAACATCATTTGAAGCCAGCAGGAGGAAAAACGACAGCTT 900
Qy 901 TTATTTTCTTATACAGCATGATCAGAACCGAGTCTGAAGATATGGGTATGCGTGAAA 960
Db 901 TTATTTTCTTATACAGCATGATCAGAACCGAGTCTGAAGATATGGGTATGCGTGAAA 960
Qy 961 GGAGATCACACCGATTTATCTGTAGTCAATGGCGATGCCATATAAACCGTCTATT 1020
Db 961 GGAGATCACACCGATTTATCTGTAGTCAATGGCGATGCCATATAAACCGTCTATT 1020
Qy 1021 GCTGAAAACTGACGGGCAAGGGAATTTATGCTCTCTGCAATTCGGCCGCCAACCGTTGCG 1080
Db 1021 GCTGAAAACTGACGGGCAAGGGAATTTATGCTCTCTGCAATTCGGCCGCCAACCGTTGCG 1080
Qy 1081 CCGGTTGAAGCCGGATTGGAAGCTTGGGACGACAGTCCAGATCAGGGAATGAGTTTATA 1140
Db 1081 CCGGTTGAAGCCGGATTGGAAGCTTGGGACGACAGTCCAGATCAGGGAATGAGTTTATA 1140
Qy 1141 AAATAAAAAAGCACCTGAAAAGGTGCTTTTTTTTGGTGTGTTTGAACCTGCTCTTCTT 1200
Db 1141 AAATAAAAAAGCACCTGAAAAGGTGCTTTTTTTTGGTGTGTTTGAACCTGCTCTTCTT 1200
Qy 1201 ATCTGTATACATAGAAATAACGTCATTTTTTATTTTTTATTTAGTCTGCTGAAAAGGTGCG 1260
Db 1201 ATCTGTATACATAGAAATAACGTCATTTTTTATTTTTTATTTTATTTAGTCTGCTGAAAAGGTGCG 1260
Qy 1261 TTGAAGTGTGGTATGTTGTTTTTAAAGTATTGAAACCTTTAAATTTGGTTGTCACAG 1320
Db 1261 TTGAAGTGTGGTATGTTGTTTTTAAAGTATTGAAACCTTTAAATTTGGTTGTCACAG 1320
Qy 1321 AAAAAACCCATCTGTATAAGTTATAAGTGACTAAACAAATAACTAAATAGATGGGGTTT 1380
Db 1321 AAAAAACCCATCTGTATAAGTTATAAGTGACTAAACAAATAACTAAATAGATGGGGTTT 1380
Qy 1381 CTTTAAATATTATGTGCTCTTAATAGTAGCATTTTATTCAGATGAAAAATCAAGGGTTTTAG 1440
Db 1381 CTTTAAATATTATGTGCTCTTAATAGTAGCATTTTATTCAGATGAAAAATCAAGGGTTTTAG 1440

Qy 1441 TGGACAAGACAAAAGTGGBAAGTGAGACCATGCTTAGGAAGACGAGTTATTAATAG 1500
Db 1441 TGGACAAGACAAAAGTGGBAAGTGAGACCATGCTTAGGAAGACGAGTTATTAATAG 1500
Qy 1501 CTGAATAAGAACGGTGTCTCCAAATATTCTTATTTAGAAAAGCAAAATCTAAAAATTAATCT 1560
Db 1501 CTGAATAAGAACGGTGTCTCCAAATATTCTTATTTAGAAAAGCAAAATCTAAAAATTAATCT 1560
Qy 1561 GAAAAAGGAATGAGAAATAGTGAATAGGACCAATAATAATAGCTAGAGAAGAAGAAATGAAG 1620
Db 1561 GAAAAAGGAATGAGAAATAGTGAATAGGACCAATAATAATAGCTAGAGAAGAAGAAATGAAG 1620
Qy 1621 ATTGTTTCATGAATAATAGGAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATT 1680
Db 1621 ATTGTTTCATGAATAATAGGAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATT 1680
Qy 1681 GGTGTTTATGGCTCTCTTGGTCTGTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATG 1740
Db 1681 GGTGTTTATGGCTCTCTTGGTCTGTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATG 1740
Qy 1741 TGTGTCATGTCACACAGAGGAAGCAGAGTTTACGCCATGAATGGACAACCCGGTGAAGTGAAG 1800
Db 1741 TGTGTCATGTCACACAGAGGAAGCAGAGTTTACGCCATGAATGGACAACCCGGTGAAGTGAAG 1800
Qy 1801 GTGGAAGTGAATTTTGTATAGCGAAGAGATTTCTACTAGATTATGTCATCTCAGGTGGAATCA 1860
Db 1801 GTGGAAGTGAATTTTGTATAGCGAAGAGATTTCTACTAGATTATGTCATCTCAGGTGGAATCA 1860
Qy 1861 GATTGGCCGGCTTACACATGTGTCATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGA 1920
Db 1861 GATTGGCCGGCTTACACATGTGTCATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGA 1920
Qy 1921 TACTTAGAAGAAAGTATCAAACTGCTAAATCGGTAGAGGCCAAACGTTCCACGATGCG 1980
Db 1921 TACTTAGAAGAAAGTATCAAACTGCTAAATCGGTAGAGGCCAAACGTTCCACGATGCG 1980
Qy 1981 ATTTGTGCGCTTATCGGTAGAGAGCTGTTTGAATATGCAGGCAAAATGCGGTAATATTCGT 2040
Db 1981 ATTTGTGCGCTTATCGGTAGAGAGCTGTTTGAATATGCAGGCAAAATGCGGTAATATTCGT 2040
Qy 2041 GTGCAAGACGCGCAACAACTTTTACCATCTCTGATCTGTACAGGTAGCAATGGCAGGTGCC 2100
Db 2041 GTGCAAGACGCGCAACAACTTTTACCATCTCTGATCTGTACAGGTAGCAATGGCAGGTGCC 2100
Qy 2101 ATGTTGATGTTGTTGTCATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
Db 2101 ATGTTGATGTTGTTGTCATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
Qy 2161 GCAGTTAAGCAATCAGATCTTCTTACGAGTTATGACCATCTGTGCCAGTTCGTAAATGTC 2220
Db 2161 GCAGTTAAGCAATCAGATCTTCTTACGAGTTATGACCATCTGTGCCAGTTCGTAAATGTC 2220
Qy 2221 GGTCAACTTTCCGACTCTGAGAAAATTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280
Db 2221 GGTCAACTTTCCGACTCTGAGAAAATTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280
Qy 2281 CAGAGTGGACAGAAACGACACGGATATATAGTGAATGTCGTCAAAACGACATACCAATTTTGA 2340
Db 2281 CAGAGTGGACAGAAACGACACGGATATATAGTGAATGTCGTCAAAACGACATACCAATTTTGA 2340
Qy 2341 ATTGGAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGGAAACGCTGCTTGAAGCG 2400
Db 2341 ATTGGAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGGAAACGCTGCTTGAAGCG 2400
Qy 2401 CGAAGCGGGCGCAGATCTGAAATATCGGCACATATTTGTTATCGTGGCAAGCGGACAGGTC 2460
Db 2401 CGAAGCGGGCGCAGATCTGAAATATCGGCACATATTTGTTATCGTGGCAAGCGGACAGGTC 2460
Qy 2461 CGTCTAACAGAGAAAGTGGATCAGGTCTGATGCGGTTTACGGAATTTAAAGACGCTATG 2520
Db 2461 CGTCTAACAGAGAAAGTGGATCAGGTCTGATGCGGTTTACGGAATTTAAAGACGCTATG 2520

Qy 722 CCTTAAGCAAAAGCTGTTGGCGGGAAGAGGTTTTTGGCGAGGATCAGCGGTCTTCATCG 781
Db CTTTAAAGCAAAAGCTGTTGGCGGGAAGAGGTTTTTGGCGAGGATCAGCGGTCTTCATCG 3303
Qy 782 ACTTTTGTCTGAACCATGCCAGAACATTTATCTTTCAAACCGCTATTCCGCCAGCCAGCT 841
Db ACTTTTGTCTGAACCATGCCAGAACATTTATCTTTCAAACCGCTATTCCGCCAGCCAGCT 3363
Qy 842 GTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGGGAAGAACGACAGCTTT 901
Db GTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGGGAAGAACGACAGCTTT 3423
Qy 902 TATTTTCTTATATCAGCATGATCAGAACCACTCTGAAGAATATGGGTTATCTGTGGAAG 961
Db TATTTTCTTATATCAGCATGATCAGAACCACTCTGAAGAATATGGGTTATCTGTGGAAG 3483
Qy 962 GAGATCACACACCGATTAATCTGTAGTCAATTTGGCGATGCCATAAACAACGCTCTCTATTG 1021
Db GAGATCACACACCGATTAATCTGTAGTCAATTTGGCGATGCCATAAACAACGCTCTCTATTG 3543
Qy 1022 CTGAATAACTCGAGGCAAGGGAATTTATGCTCTGCCATTCGGCCGCCAACCGTTGGC 1081
Db CTGAATAACTCGAGGCAAGGGAATTTATGCTCTGCCATTCGGCCGCCAACCGTTGGC 3603
Qy 1082 CGGGTGAAGCCGGATTCGAAGCTTGGGCAGCAGTTCGAGATCAGGGGAATGAGTTTATAA 1141
Db CGGGTGAAGCCGGATTCGA-----ATTA 3627
Qy 1142 AATAAAAAAGCAGCTGAAGAGTGCTTTTTTTTGTATGGTTTTGAACTGTTCTTTCTTA 1201
Db CAATCACGCTCTGACCAACAGTAGTAGGTGATATGATCATTTGCTGCAAAACATTTCAITCAA 3687
Qy 1202 TCTTGATACATATAGNAATACGTCATTTTATTTTATTTTATTTTGTCTGGAAGGTGGT 1261
Db TCGAAGAGGAGCTGCACATCATTTGAGGGGTTTTTTGTGACGGGAACCTGA----- 3738
Qy 1262 TGAAGTGTGGTATGATGTGTTTAAAGTATGAAAACCCCTTAAATTTGGTTGCACAGA 1321
Db -----TACAGAAGTAGGGAAGAACGGTTATATCCAGCGTCTTG- 3776
Qy 1322 AAAACCCCATCTGTTAAAGTTTAAAGTGAATAAACAATAACTAAATAGATGGGGGTTTC 1381
Db -----CTGGCTTATTGAAGACAATAATAGACATGTGCGGGGTGA 3816
Qy 1382 TTTTAAATATTATGTCTCTAATAGTACATTTATTCAGATGAATAAATAAAGGTTTATG 1441
Db TAAACCATTTTAAAGCGGGATATCGGCCATCATCCAGATAGTATACAAAGTTTGTCTGAA 3876
Qy 1442 GGACAAGACAAAAAGTGGAAAAGTGAGACCATGTGCTTAGGAAGACGAGTTATTAAATAGC 1501
Db AGATATGTCGAGA----- 3890
Qy 1502 TGAATAAGACGGTGCTCTCAAATATTTCTTATTTAGAAAAGCAAAATCTAAAATTAATCTG 1561
Db -----CCAGTCTTCTCATGAAGACATTACGCCCTTTGCGCTTCAA 3930
Qy 1562 AAAAGGGAATGAGAAATAGTGAATGGAACCAATAATATGACTAGAGAAGAAAGATGAAGA 1621
Db GGCCCGCTTGCAACCATACGTTGCGAGGAAACTTTGAGGGAAAAGACTGTACCATCTGGAAGA 3990
Qy 1622 TTGTTTCATGAAATTAAGGAACGAATATTTGGATAAATATGGGGATGATGTTAAGGCTATTG 1681
Db GGTTTTAAAGCCATTTGGGGCGGATAGAGAAAA-----CATGAATGCTTCATCGTAG 4043
Qy 1682 GTGTTTATGGCTCTCTTGGTCGACAGTGAAGGGCCCTATTTCGGATATTGAGATGATGT 1741
Db AAGGTGACGGCGGTATTCTGTGCCATTTGGGAGAGGACTATTTGGTCAGTCATGTCATAA 4103
Qy 1742 GTGTCATGTCACAGAGGAGACAGATTCACCCATGAATGGAACAACCGGTGAGTGGGAAG 1801
Db -----AAGCGTTGACGCTTCCCATGATTAATTTGTGGCGGCTCTCGCCTTGGAAACC 4153
Qy 1802 TGGAAAGTGAATTTTGTATAGCAAGAGATTTCTACTAGATTTATGCAATTCAGGTGGAATCAG 1861

Db 4154 ATTAATCATACCTTTTTTAACTGTCAAATATGCAAGAAAGCATGGGGCTTCCAATCGCCGA 4213
Qy 1862 ATTGGCCCGCTTACACATGGTCAATTTTTTCTCTATTTTGCCGATTTATGATTCAAGTGGAT 1921
Db 4214 ATTATCATCAATGGAATCAGTGAATCTCTCTGATGAAGATGAAAAAACCAATCTCTGAGATG 4273
Qy 1922 ACTTAGAGAAAAGTGTATCAAACTGCTAAATCGGTAGAGGCCAAACGTTGCCACATGCGA 1981
Db 4274 ATTGAGCGCTTATGCGGTGTCGCGATTTTAGGGGTTAGCCAAAGCTTGCACACGTG--- 4330
Qy 1982 TTTGTGCCCTTATCTGTAGAAAGAGCTGTTTGAATATGAGGCAAAATGCGGTAAATATTCGTG 2041
Db 4331 -----ACGAAGAAAACGGTTCTACATATGTTAAAGACCAT----- 4367
Qy 2042 TGCAGGACCGACAAACATTTCTACCATCTTGACTGTACAGGTAGCAATGGCAGGTGCGCA 2101
Db 4368 -----TCNAATCTATTACTGATGAATCAAGTGGGGGTATGA 4405
Qy 2102 TGTGTATTGCTGTCATCATCGCATCTGTTTATACGACGAGCGTTCGGTCTTAACCTGAAG 2161
Db 4406 GAATGAATCAATGATGGAACCTGCAGACCGGGTGTCTGGCTGGAGCAGAAAGTGACTGACG 4465
Qy 2162 CAGTTAAGCAATCAGATCTTCTCTTACAGGTTATGACCATCTGTGCCAGTTCGTAAATGCTG 2221
Db 4466 AAGAGGCGCTTTCAATATTACATTTGCTCTGATGAAGATATTTTCTGCTATTAATGCACGGG 4525
Qy 2222 GTCAACTTTCGGACTCTGAGAAACTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTC 2281
Db 4526 CTTTTTCATCAGAAAAACACTTTTACGGAANAAGTAAGTCAANTATGATTATGAATG 4585
Qy 2282 AGAGTGGACAGAAACGACGAGATATATAGTGGATGTCAAAAACGCATACCAATTTTGA 2341
Db 4586 CGAAATCCGGGCTCTGCCCGAAAACTGCGGCTATTGTTTACA-----GTCTGCGAT 4637
Qy 2342 TTCGAAAACGCCGATTGAGTCTTAACGGATGGTGAATAAGAAAAACGCTGCTTGAAGGCGC 2401
Db 4638 TTCGAAAACGCCGATTGAGTCTTACCGGATGGTGAATAAGAAAAACGCTGCTTGAAGGCGC 4697
Qy 2402 GAAGCGGGCGCACGATCTGAATATCGGCACATATTGATCGTGGCAGCGCAGAGTCC 2461
Db 4698 GAAGCGGGCGCACGATCTGAATATCGGCACATATTGATCGTGGCAGCGCAGAGTCC 4757
Qy 2462 GTCTAACAGAGAAGTGGATCAGGTCTGATGCGGTTTCAAGAAATTAAGAGACGTATGG 2521
Db 4758 GTCTAACAGAGAAGTGGATCAGGTCTGATGATGCGGTTTCAAGAAATTAAGAGACGTATGG 4817
Qy 2522 ACTGAAGATTGTGCAATGCTTTGGACTGTTGAAGCCAGACAGCGGAGCGGCTCAAGA 2581
Db 4818 ACTGAAGATTGTGCAATGCTTTGGACTGTTGAAGCCAGACAGCGGAGCGGCTCAAGA 4877
Qy 2582 TGCAGGATGACCGCTTATTAATCATTAATTTGAATACGTCAAGAGAAACCAATTTCAACAT 2641
Db 4878 TGCAGGATGACCGCTTATTAATCATTAATTTGAATACGTCAAGAGAAACCAATTTCAACAT 4937
Qy 2642 CACAACCTCACATACATACGATGACAGATCAATACGTTTGAATTCGCAAAAGAAATCGGG 2701
Db 4938 CACAACCTCACATACATACGATGACAGATCAATACGTTTGAATTCGCAAAAGAAATCGGG 4997
Qy 2702 GCTGTCTCCGTGTTCAAGCGCCATTATCGGATGAAGAGACGAAACAGGATGCTATGA 2761
Db 4998 GCTGTCTCCGTGTTCAAGCGCCATTATCGGATGAAGAGACGAAACAGGATGCTATGA 5057
Qy 2762 CATCGCCAAAAGCTTGAAGGCTCTTGACCGGGATTCATTTCTGTGAATTTTTTGCATGC 2821
Db 5058 CATCGCCAAAAGCTTGAAGGCTCTTGACCGGGATTCATTTCTGTGAATTTTTTGCATGC 5117
Qy 2822 AATTGATGGCAGCGCGTTAGAAAGCGCTCAACGAATTAACCCGCTGTATTGTTTAAAGT 2881
Db 5118 AATTGATGGCAGCGCGTTAGAAAGCGCTCAACGAATTAACCCGCTGTATTGTTTAAAGT 5177
Qy 2882 GCTGGCGCTGTTCCGTTTATCAATCCATCAAAAAGAAATTCGCAATTTCCGAGAGGAAGA 2941

Db 5178 GCTGGCGCTGTCGTTTATCAATCCATCAAAAGAAATTCGCAATTTCCGGAGGAGAGA 5237
Qy 2942 GGTCAATCTCGCACATTTGAGCCATTAGGCTTTACGCGCAAACTCCCAATTTTGTGCGG 3001
Db 5238 GGTCAATCTCGCACATTTGAGCCATTAGGCTTTACGCGCAAACTCCCAATTTTGTGCGG 5297
Qy 3002 AGACTACTTAACAACTCCCGGCGAAGAGAGAGAGAGATCATATAAAATGCTGAGTGATTT 3061
Db 5298 AGACTACTTAACAACTCCCGGCGAAGAGAGAGAGAGATCATATAAAATGCTGAGTGATTT 5357
Qy 3062 AGGCTTTGAAGTTGAATCAGTCGAGAAATGAGGCTAGTTTAAGTCGGAAGACTGAAA 3121
Db 5358 AGGCTTTGAAGTTGAATCAGTCGAGAAATGAGGCTAGTTTAAGTCGGAAGACTGAAA 5417
Qy 3122 GAATCAATAAAGCAATCGGTATGATGTCGAAT 3154
Db 5418 GAATCAATAAAGCAATCGGTATGATGTCGATT 5450

RESULT 3
US-09-407-549-1
; Sequence 1, Application US/09407549
; Patent No. 6303377
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley Grant
; APPLICANT: Perkins, John B.
; APPLICANT: Vocum, R. Rogers
; APPLICANT: Pero, Janice G.
; TITLE OF INVENTION: BIOTIN BIOSYNTHESIS IN BACILLUS
; NUMBER OF INVENTION: SUBTILIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,549
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/239,430
; FILING DATE: May 6, 1994
; APPLICATION NUMBER: 08/084,709
; FILING DATE: June 25, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,065
; REFERENCE/DOCKET NUMBER: 04599/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8478
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-407-549-1

Query Match 51.1%; Score 1612.6; DB 3; Length 8478;
Best Local Similarity 73.0%; Pred. No. 0;
Matches 2303; Conservative 0; Mismatches 624; Indels 226; Gaps 9;
Qy 2 GATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGTTAAACGAGCGGTTAGACAGAA 61

2524 GATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGAA 2583
Qy 62 TGAAGAAGCGCGCTACATCGTAACTCGCGTCAATGGATGAGCGCCGGTTCACAGAGA 121
Db 2584 TGAAGAAGCGCGCTACATCGTAACTCGCGTCAATGGATGAGCGCCGGTTCACAGAGA 2643
Qy 122 GGAATATTGATGCGGAAATCAAAACGGTCTGGTCTCTCAAAACAATTTATTAGGGCTCCAA 181
Db 2644 GGAATATTGATGCGGAAATCAAAACGGTCTGGTCTCTCAAAACAATTTATTAGGGCTCCAA 2703
Qy 182 GCATAGACGTTTATGATGATGAGCCCAACAGCATTTGAGCAATTTGGGACAGGAAGCA 241
Db 2704 GCATAGACGTTTATGATGATGAGCCCAACAGCATTTGAGCAATTTGGGACAGGAAGCA 2763
Qy 242 GCGGTTACGTTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAAAGATTG 301
Db 2764 GCGGTTACGTTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAAAGATTG 2823
Qy 302 CCAGCTTTAAACTGACAGAGCGGCTGCTGTTTTCGAGCGGTTACTTGGCCCAATGTCG 361
Db 2824 CCAGCTTTAAACTGACAGAGCGGCTGCTGTTTTCGAGCGGTTACTTGGCCCAATGTCG 2883
Qy 362 GTGTCCTTTTCATCTTCCAGAAAAGGAAAGATGTCATTTTAAGTGACAGCTCAATCATG 421
Db 2884 GTGTCCTTTTCATCTTCCAGAAAAGGAAAGATGTCATTTTAAGTGACAGCTCAATCATG 2943
Qy 422 CAAGTATGATCGACGGCTCGGACTTTCTAAGGCTGATACAGTTGTTTATCGGCATATTG 481
Db 2944 CAAGTATGATCGACGGCTCGGACTTTCTAAGGCTGATACAGTTGTTTATCGGCATATTG 3003
Qy 482 ATATGAATGATCTTTGAAAAACAAGCTGAATGAAAAACACAGCGTTATCAGCGCGGTTTTATCG 541
Db 3004 ATATGAATGATCTTTGAAAAACAAGCTGAATGAAAAACACAGCGTTATCAGCGCGGTTTTATCG 3063
Qy 542 TAACAGACGAGGATTATCAGCATGATGAGCAATTCGCCCTCTTGTATCAGATCATCATCAG 601
Db 3064 TAACAGACGAGGATTATCAGCATGATGAGCAATTCGCCCTCTTGTATCAGATCATCATCAG 3123
Qy 602 TTGCGAAACGCTATCATGCTCTGCTGCTGTTGATGATGCCAGCAACAGGAGTTTGG 661
Db 3124 TTGCGAAACGCTATCATGCTCTGCTGCTGTTGATGATGCCAGCAACAGGAGTTTGG 3183
Qy 662 GCGATTGCGGACCAAGGAAACAGTGAAATPACTTTGGTGTGTTGTCGCCGACATTGTTATCGGCA 721
Db 3184 GCGATTGCGGACCAAGGAAACAGTGAAATPACTTTGGTGTGTTGTCGCCGACATTGTTATCGGCA 3243
Qy 722 CTTTAAGCAAAAGCTTTGGCGGAGGAGGTTTGGCGGAGGATCAGCGGCTCTTCATCG 781
Db 3244 CTTTAAGCAAAAGCTTTGGCGGAGGAGGTTTGGCGGAGGATCAGCGGCTCTTCATCG 3303
Qy 782 ACTTTTGTGTAACCATGCGCAGAACATTTATCTTTCAAAACCGCTATTTCCGCCAGCCAGCT 841
Db 3304 ACTTTTGTGTAACCATGCGCAGAACATTTATCTTTCAAAACCGCTATTTCCGCCAGCCAGCT 3363
Qy 842 GTGCGGCTGCTCAGCGGCTTTTCAACATCATTTGAAGCCAGCAGGAGAAAACGACAGCTTT 901
Db 3364 GTGCGGCTGCTCAGCGGCTTTTCAACATCATTTGAAGCCAGCAGGAGAAAACGACAGCTTT 3423
Qy 902 TATTTTCTTATATCAGCATGATCAGAAACAGTCTGAAGAATATGCGGTTATGTTGTGAAG 961
Db 3424 TATTTTCTTATATCAGCATGATCAGAAACAGTCTGAAGAATATGCGGTTATGTTGTGAAG 3483
Qy 962 GAGATCACACACCGATTATTTCTGTAGTCAATGCGCATGCCCAATAAAACGGTCTCTATTG 1021
Db 3484 GAGATCACACACCGATTATTTCTGTAGTCAATGCGCATGCCCAATAAAACGGTCTCTATTG 3543
Qy 1022 CTGAAAAAATGACAGGCAAGGAAATTTATGCTCTGCTGCAATTCGCGCCGCAACCGTTGCGC 1081
Db 3544 CTGAAAAAATGACAGGCAAGGAAATTTATGCTCTGCTGCAATTCGCGCCGCAACCGTTGCGC 3603
Qy 1082 CGGTTGAACCGGATTCGAAGCTTTGGGACAGGCTCGAGATCAGGGAATGAGTTTATAA 1141

Db 3604 CGGGTGAAGCGGGATTGCA-----ATTA 3627
Qy 1142 ANTAAGAAAGCCTGAAAGGTGTCTTTTGTGATGGTTTGAACCTGTTCTTTCTTA 1201
Db 3628 CAATACGCTCGACACAGTAGTGGGTGATATGTATCATTTGCTGCAACATTTTCATTCAA 3687
Qy 1202 TCTTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTTATTTTGTCTGAAAGGTGGT 1261
Db 3688 TCGNAAGGAGCTGCACATCATTTGAGGGGTTTTTTTGTGACGGGACTGA----- 3738
Qy 1262 TGAAGTGTGTGATGTATGTGTTTAAAGTATGAAAAACCTTAAATTTGTTGACACAGA 1321
Db 3739 -----TACAGAAGTAGGGAAACCGGTATATCCAGCGGTCTTG- 3776
Qy 1322 AAAACCCCATCTGTTAAAGTTATAAGTGACATAACCAATACTAAATAGATGGGGTTTC 1381
Db 3777 -----CTGCGCTTATTTGAAAGACAATAATAGACATGTTCGGGGTGA 3816
Qy 1382 TTTTAAATATTGTCTCTAATAGTAGCATTTATTCAGATGAAAAAATCAAGGGTTTTAGT 1441
Db 3817 TAAACCATTTTAAAGCGGATATCGGCCATCATCCAGATAGTATACAGTTTGTCTGAA 3876
Qy 1442 GGACAAGCAAAAAGTGGAAAAGTAGACCATGTGCTTTAGGAAGACGAGTTATTAAATAGC 1501
Db 3877 AGATATGTGCGAGA----- 3890
Qy 1502 TGAATAAGAACGGTCTCTCCAAATATCTTATTTAGAAAAGCAAACTTAAATTTACTG 1561
Db 3891 -----CGAGCTTTCTCATGAAGACATATCGCCTTTTGCCCTTCAA 3930
Qy 1562 AAAAGGGAATGAGAATAGTAGAATGACCAATAATAATGACTAGAGAAGAAAGATGAAGA 1621
Db 3931 GCGCGCGCTTGCAACATACGTTGACGGGAACCTGAGGGAAGACTGTCACCATGGAAGA 3990
Qy 1622 TTGTTTCAAGAAATTAAGAAAGATATTTGGATAAATATGGGGATGATTTAAGGCTATTG 1681
Db 3991 GGTTTTAAAGCCATTTGGGGCGGATTTAGAGAAAA-----CATGAATGCTTCATCGTAG 4043
Qy 1682 GTGTTTATGSGCTCTCTGGTCGTAGACTGATGGCCCTATTCGGATATTAGATGATGT 1741
Db 4044 AAGTGCAGGCGGATTTCTGTGCCATTTGGAGAGGACTATTTGGTCAGTCATGTCTATAA 4103
Qy 1742 GTGTCATGTCAAAGAGGAAGCAGAGTTTCAGCCCATGAATGACAAACCGGTGAGTGGGAAG 1801
Db 4104 -----NAGCGTGCAGCTTCCCATGATTATTTGGGCGCTCTCGCTTGGAAACC 4153
Qy 1802 TGAAGTGAATTTGTATAGGAAGAGATTTCTAGATATATGATCTCAGGTGGAATCAG 1861
Db 4154 ATTAATCATACCTTTTAACTGTCAAATATATGCAAGAAAGCATGGGCTTCCCAATGCCGGA 4213
Qy 1862 ATTGGCCGCTTACATGGTCAATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGAT 1921
Db 4214 ATTATCATCAATGGAATCAGTGACTCTCTGATGAAGATGAAGAAACCAATCTCTGAGATG 4273
Qy 1922 ACTTAGAGAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAACGTTCCACCATGCGA 1981
Db 4274 ATTGAGCGCTTATGGGGTGCAGATTTTAGGGGTACGCCAAAGCTTGCACACGTG--- 4330
Qy 1982 TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGAGGCAATATGGCGTAAATATTCGTG 2041
Db 4331 -----ACGAAAGAAACCGGTTCTACATATGGTAAAGACCAATA----- 4367
Qy 2042 TGCAAGGACCACACATCTTACCATCTTGTAGCTGACAGTAGCAATGCGAGGTGCCA 2101
Db 4368 -----TCAATCTATCATTACTGATGAATCAAGTGGGGGTATGA 4405
Qy 2102 TGTGAATGGTCTGCATCATCGCATCTGTTATAGCAGAGCGCTTCGGTCTTTAACTGGAAG 2161
Db 4406 GAATGAATCAATGATGGAATCTGCAGACCGGGTGTGGTGGAGCAGAGTACTGACG 4465
Qy 2162 CAGTTAAGCAATCAGATCTTCTCAGGGTTAAGACCAATCTGTGCGAGTTCGTGAATGCTG 2221
Db 4466 AAGAGGCGCTTTCAATATTAATTTGTCCTGATGAAGATATTTTGTCTATTAATGACCGGG 4525

Qy 2222 GTCAACCTTTCCGACTCTGAGAAACCTTCTGGAATCGCTAGAGAAATTTCTGGAATGGATT 2281
Db 4526 CTTTTTCATCAGAAAAACCTTTTACGAAAAAAAAGTAAGCTCAATATGATTATGATG 4585
Qy 2282 AGGAGTGGACAGAACGACACGGATATATAGTGGATGTGTCAAAACGCTATACCATTTTGAA 2341
Db 4586 CGAAATCCGGGCTCTGCCCGGAAAACTGCGGCTATTGTTTCA-----GTCTGCGAT 4637
Qy 2342 TTTGAAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGGAAAACGCTGCTTGAAGGCG 2401
Db 4638 TTTGAAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGGAAAACGCTGCTTGAAGGCG 4697
Qy 2402 GAAGCGGCGCACGATCTGGAATATCGGCACATATTTGATCGTGGCAAGCGCAGAGGTCC 2461
Db 4698 GAAGCGGCGCACGATCTGGAATATCGGCACATATTTGATCGTGGCAAGCGCAGAGGTCC 4757
Qy 2462 GTCTAAACAGAGAAGTGGATCAGGTCTGATGTCGGTTTCAAGAAATTTAAAGAGAGCTATGG 2521
Db 4758 GTCTAAACAGAGAAGTGGATCAGGTCTGATGTCGGTTTCAAGAAATTTAAAGAGAGCTATGG 4817
Qy 2522 ACTGAAGATTTGTGATGCTTTGAGCTGTTGAAGCCAGAGCAGCGGAAGCGGCTCAAGA 2581
Db 4818 ACTGAAGATTTGTGATGCTTTGAGCTGTTGAAGCCAGAGCAGCGGAAGCGGCTCAAGA 4877
Qy 2582 TGCAAGAGTAGACCGCTATTAATCATTAATTTGAATACGTACAGAGAAACCATTTCAAACAT 2641
Db 4878 TGCAAGAGTAGACCGCTATTAATCATTAATTTGAATACGTACAGAGAAACCATTTCAAACAT 4937
Qy 2642 CACAACCTCACAATACATACATGAGTCAAGTCAATACCGTTGAAATCGCAAAAGAAATCGGG 2701
Db 4938 CACAACCTCACAATACATACATGAGTCAAGTCAATACCGTTGAAATCGCAAAAGAAATCGGG 4997
Qy 2702 GCTGTCTCGTGTTCAGGCGCCATTTATCGGATGAAGAGACGAAACAGGATGTCATTGA 2761
Db 4998 GCTGTCTCGTGTTCAGGCGCCATTTATCGGATGAAGAGACGAAACAGGATGTCATTGA 5057
Qy 2762 CATCGCCAAAGCTTTGAAGGCTCTTGACGCGGATTCATTCCTGTAATTTTTTGTGATGC 2821
Db 5058 CATCGCCAAAGCTTTGAAGGCTCTTGACGCGGATTCATTCCTGTAATTTTTTGTGATGC 5117
Qy 2822 AATTGATGGCAGCGCTTTAGAAGCGTCAACGAATTTAAACCCGCTGTATTGTTTAAAGT 2881
Db 5118 AATTGATGGCAGCGCTTTAGAAGCGTCAACGAATTTAAACCCGCTGTATTGTTTAAAGT 5177
Qy 2882 GCTGGCGCTGTTCGGTTTTATCAATCCATCAAAAGAAATTTCCGATTTCCGGAGGAAGA 2941
Db 5178 GCTGGCGCTGTTCGGTTTTATCAATCCATCAAAAGAAATTTCCGATTTCCGGAGGAAGA 5237
Qy 2942 GGTCAATCTCGCACATTTGACGCCATTTAGGCTTTAGCGCGCAAACTCCATTTTGTGCG 3001
Db 5238 GGTCAATCTCGCACATTTGACGCCATTTAGGCTTTAGCGCGCAAACTCCATTTTGTGCG 5297
Qy 3002 AGACTACTTAACAACTCCCGGCAAGAGAGACGAGGATCATATAAATGCTGAGTGATTT 3061
Db 5298 AGACTACTTAACAACTCCCGGCAAGAGAGAGAGAGAGAGATCATATAAATGCTGAGTGATTT 5357
Qy 3062 AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTGGAAAAAGCTGAAA 3121
Db 5358 AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTGGAAAAAGCTGAAA 5417
Qy 3122 GAATCAATAAAGCAATCGGTATGATGTCGAAT 3154
Db 5418 GAATCAATAAAGCAATCGGTATGATGTCGAAT 5450

RESULT 4

US-09-407-728-1
; Sequence 1, Application US/09407728
; Patent No. 6841366
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley Grant
; APPLICANT: Perkins, John B.

APPLICANT: Yocum, R. Rogers
 APPLICANT: Pero, Janice G.
 TITLE OF INVENTION: BIOTIN BIOSYNTHESIS IN BACILLUS
 TITLE OF INVENTION: SUBTILIS
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:

Query Match	51.1%	Score 1612.6;	DB 3;	Length 8478;
Best Local Similarity	73.0%	Pred. No. 0;		
Matches 2303;	Conservative	0;	Mismatches 624;	Indels 226;
Gaps 9;				Gaps 9;

Qy	2	GATCCACGAGGTTACGAGCCTTGAAGATATGATTCCTGGTTAAACGAGCGGTTAGACAGAA	61
Db	2524	GATCCACGAGGTTACGAGCCTTGAAGATATGATTCCTGGTTAAACGAGCGGTTAGACAGAA	2583
Qy	62	TGAAGAAGCGGGGTACATCGTAAACCTGCGGTCAATGGATGGAGCGCCGGTTCACAGA	121
Db	2584	TGAAGAAGCGGGGTACATCGTAAACCTGCGGTCAATGGATGGAGCGCCGGTTCACAGA	2643
Qy	122	GGAATATTGATGGCGAAATCAAAACGGTCTGGTCTCTCAAAACAATTAATTAGGGGCTCGCAA	181
Db	2644	GGAATATTGATGGCGAAATCAAAACGGTCTGGTCTCTCAAAACAATTAATTAGGGGCTCGCAA	2703
Qy	182	GCGATAGACGTTTGTATCGATTCGAGCCCAAAACAGCATTTGCAGCAATTTGGGACAGAAAGCA	241
Db	2704	GCGATAGACGTTTGTATCGATTCGAGCCCAAAACAGCATTTGCAGCAATTTGGGACAGAAAGCA	2763
Qy	242	CGGGTTACGTTTAAACGACAGGCNAATTGCGTCTGGCATGAAGCTAGAAAGAGAGATTG	301
Db	2764	CGGGTTACGTTTAAACGACAGGCNAATTGCGTCTGGCATGAAGCTAGAAAGAGAGATTG	2823
Qy	302	CCAGCTTTAAACTGACAGAAACGGCCCTCTGTTTTCGAGCGGTTACTTTGGCCCAATGTCG	361
Db	2824	CCAGCTTTAAACTGACAGAAACGGCCCTCTGTTTTCGAGCGGTTACTTTGGCCCAATGTCG	2883
Qy	362	GTGTCTCTTTTCATCTTCCCGAAGAAAGATGTCATTTTTAAGTGAACAGCTCAATCATG	421
Db	2884	GTGTCTCTTTTCATCTTCCCGAAGAAAGATGTCATTTTTAAGTGAACAGCTCAATCATG	2943

QY 1502 TGAATAAGACGGTCTCTCCAAATATTTCTTATTAGAAAAACCAATCTAAATATATCTG 1561
Db 3891 -----CCAGTCTTTCTCATGAGACATTACGCTTTTGCCCTTCAA 3930
QY 1562 AAAAGGGAAATGAGATAGTGAATGGACCAATAATAATGACTAGAGAGAAAGAAATGAAGA 1621
Db 3931 GCGCGCGCTTGCAACCATACATAGTTGACGGGAAACTTGAGGAAAGACTGTCACCATGGAAGA 3990
QY 1622 TTGTTTCAATTAAGAAAGCAATATTTGGATAAATATGGGGATGATGTTAAGGCTATTG 1681
Db 3991 GGTTTTAAGCCATTGGGGCGGATTTAGAGAAAA-----CATGAATGCTTCATCGTAG 4043
QY 1682 GTGTTTATGGCTCTCTGGTGGTCAAGACTGATGGCCCTATTTCGGATATTCAGATGATGT 1741
Db 4044 AAGGTGACGGCGGTATTTCTGTGCCATTGGGAGGAGACTATTTGGTCAATGTCATTA 4103
QY 1742 GTGTCAATGTCAACAGAGAAAGCAGAGTTCAAGCCATGAATGACCAACCGGTGAGTGGAGG 1801
Db 4104 -----AAGCGTTGCAGCTTCCCATGATTTATTTGGCGCGCTCTCGCCTTTGGAACC 4153
QY 1802 TGGAAAGTGAATTTGATAGCAAGAGATTTACTAGATTATGATCTCAGGTGGATCAG 1861
Db 4154 ATTAATCATACCTTTTAACTGTCAATAATATGCAAGAAAGCATGGGCTTCCAAATCGCCGGA 4213
QY 1862 ATTGGCCGCTTACACATGGTCAAATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGAT 1921
Db 4214 ATTATCATCAATGNAATCAGTGACTCTCTGATGAGATGAATAAACAATCTCGAGATG 4273
QY 1922 ACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAACCCAAACGTTCCACGATGCGA 1981
Db 4274 ATTGAGCGCTTATGCGGTGTCGATTTTAGGGGTTACGCCAAAGCTTGGCCAAAGCTG-- 4330
QY 1982 TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGACGGCAAAATGGCGTAAATTTCTGTG 2041
Db 4331 -----ACGAAGAAACCGGTTTACATATGGTAAAGACCCATA----- 4367
QY 2042 TGCAAGGACCGACAACATTTTACCATCTTTGACTGTACAGGTAGCAATGCGAGTGCCTCA 2101
Db 4368 -----TCAATCTATCTACTGATGAATCAAGTGGGGTATGA 4405
QY 2102 TGTGATTTGCTGCAATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTAGAG 2161
Db 4406 GAATGAATCAATGATGTAATCGCAGACCGGGTCTGGCTGGAGCAGAAAGTCACTGACG 4465
QY 2162 CAGTTAGCAATCAGATCTTCTCAGGTTATGACCATCTGTCGACGTTCCGTAATGCTG 2221
Db 4466 AAGAGCGCTTTCAATATTAATTTCTGCTGATGAAGATTTTGTCTATTAATGACCGGG 4525
QY 2222 GTCAACTTTCCGACTCTGAGAACTTCTGGAAATCGCTAGAGAAATTTCTGGAAATGGGATTC 2281
Db 4526 CTTTTCACATCAGAAACACTTTTACGGAAAAAAGTAAAGCTCAATATGATTATGAATG 4585
QY 2282 AGGAGTGACAGAACGACACCGATATATAGTGGATGTGTCAAAACGCAATACCATTTTGAA 2341
Db 4586 CGAAATCGGGCTCTGCCCGGAAAACTCGGCTATTTGTTTCA-----GTCTGGCAT 4637
QY 2342 TCCGAAAGCGCGATTTGATCTTACCGATGTTGATGAATGAAGAAACGCTGCTTGAAGCGC 2401
Db 4638 TTCGAAAGCGCGATTTGATCTTACCGGATGTTGATGAATGAAGAAACGCTGCTTGAAGCGC 4697
QY 2402 GAAGCGGCGCACGATCTGAATATCGGCACATATTTGATCTGTGGCAAGCGCAGAGGTCC 2461
Db 4698 GAAGCGGCGCACGATCTGAATATCGGCACATATTTGATCTGTGGCAAGCGCAGAGGTCC 4757
QY 2462 GTCTAAACAGAGAGTGGATCAGGTGATAGTCCGGTTCCAGAAATTAAGAGACGATATGG 2521
Db 4758 GTCTAAACAGAGAGTGGATCAGGTGATAGTCCGGTTCCAGAAATTAAGAGACGATATGG 4817
QY 2522 ACTGAAGTTTGTGATGCTTGTGACTGTTTGAAGCCAGAGCAGCGGAGCGGCTCAAGA 2581
Db 4818 ACTGAAGTTTGTGATGCTTGTGACTGTTTGAAGCCAGAGCAGCGGAGCGGCTCAAGA 4877
QY 2582 TGCAGGAGTAGACCGCTATATATCATATTTTGAATACGTCACAGAGAAACCATTTCAAACAT 2641

Db 4878 TGCAGGAGTAGACCGCTATATATCATATTTGAATACGTCAAGAGAAACCATTTCAAACAT 4937
QY 2642 CACAACTCTCATATACATACGATGACAGAGTCAATACCGTTGAAATCGCAAAAGATCGGG 2701
Db 4938 CACAACTCTCATATACATACGATGACAGAGTCAATACCGTTGAAATCGCAAAAGATCGGG 4997
QY 2702 GCTGTCTCCGTGTTTCAGCGCGCATTTATCGGGATGAAGAGACGAAACAGGATGTCATTGA 2761
Db 4998 GCTGTCTCCGTGTTTCAGCGCGCATTTATCGGGATGAAGAGACGAAACAGGATGTCATTGA 5057
QY 2762 CATGCCAAAGCTTTGAAGGCTCTTGACGCGGATTTCCATCTCTGTGAATTTTTTGCATGC 2821
Db 5058 CATGCCAAAGCTTTGAAGGCTCTTGACGCGGATTTCCATCTCTGTGAATTTTTTGCATGC 5117
QY 2822 AATTGATGGCAGCGCTTTAGAGGCTCAAGGCTCAAGAAATTAACCCGCTGTATTTGTTAAAGT 2881
Db 5118 AATTGATGGCAGCGCTTTAGAGGCTCAAGGCTCAAGAAATTAACCCGCTGTATTTGTTAAAGT 5177
QY 2882 GCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAGAGAGA 2941
Db 5178 GCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAGAGAGA 5237
QY 2942 GGTCAATCTCCGACATTTGACGCGCATTTAGGCGCTTTACGCCGCAAACTCCATTTTGTGCG 3001
Db 5238 GGTCAATCTCCGACATTTGACGCGCATTTAGGCGCTTTACGCCGCAAACTCCATTTTGTGCG 5297
QY 3002 AGACTACTTAACTTCCGCGGCAAGAGGAGAGAGGAGATCAATAATGCTGAGTGATTT 3061
Db 5298 AGACTACTTAACTTCCGCGGCAAGAGGAGAGAGGAGATCAATAATGCTGAGTGATTT 5357
QY 3062 AGGCTTTGAAGTTGAATCAGTTCGAAGAAATGAAGGCTAGTTTAAAGTGCRAAAAGCTGAAA 3121
Db 5358 AGGCTTTGAAGTTGAATCAGTTCGAAGAAATGAAGGCTAGTTTAAAGTGCRAAAAGCTGAAA 5417
QY 3122 GAATCAATAAAGCAATCGGTATGATGTCGAAT 3154
Db 5418 GAATCAATAAAGCAATCGGTATGATGTCGAAT 5450

RESULT 5
US-09-869-855A-3
; Sequence 3, Application US/09869855A
; Patent No. 6762040
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Method For Increasing Gene Copy Number
; FILE REFERENCE: 10028.204-US
; CURRENT APPLICATION NUMBER: US/09/869,855A
; CURRENT FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5793
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE: Synthetic
; OTHER INFORMATION: Synthetic
US-09-869-855A-3

Query Match 27.5%; Score 867; DB 3; Length 5793;
Best Local Similarity 100.0%; Pred. No. 2.3e-237;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTTAT 1534
Db 968 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTTAT 1027
QY 1535 TTAGAAAAGCAATCTAAATTTATCTGAAAAGGGAATGAGATAGTGAATGACCAATAA 1594
Db 1028 TTAGAAAAGCAATCTAAATTTATCTGAAAAGGGAATGAGATAGTGAATGACCAATAA 1087
QY 1595 TAATGACTAGAGAAAGAAATGAAGATTGTTTCATGAATTTAAGGAACGAAATATTGGATA 1654

[illegible]


```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,011A
; FILING DATE: 17-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/418,085
; FILING DATE: 06-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/054,185
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/002,608
; FILING DATE: 11-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/474,857
; FILING DATE: 30-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/474,798
; FILING DATE: 16-JULY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL89/00072
; FILING DATE: 25-SEPT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL88/200904.6
; FILING DATE: 06-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL/88/202080.3
; FILING DATE: 03-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146.1169-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7336 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
; FEATURE:
; OTHER INFORMATION: PLASMID pBHA-1
; US-09-099-011A-3

Query Match 27.5%; Score 867; DB 3; Length 7336;
Best Local Similarity 100.0%; Pred. No. 2.6e-237;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGACGAGTATTATAGCTGAATAAGACGGTGTCTCTCAATATTCTTAT 1534
DB 5231 TGCTTAGGACGAGTATTATAGCTGAATAAGACGGTGTCTCTCAATATTCTTAT 5290
QY 1535 TTAGAAAAGCAATCTAAAATTATCTGAAAAGGAATGAGATAGTGAATGGACCAATAA 1594
DB 5291 TTAGAAAAGCAATCTAAAATTATCTGAAAAGGAATGAGATAGTGAATGGACCAATAA 5350
QY 1595 TAATGACTAGAGAAGAAGATGAAGATTTTCATGAAATTAAGGAACGAATATTGGATA 1654
DB 5351 TAATGACTAGAGAAGAAGATGAAGATTTTCATGAAATTAAGGAACGAATATTGGATA 5410
QY 1655 AATATGGGATGATGTTAAGGCTATTGGTCTTTATGCTCTCTTGGTCGTCAGACTGATG 1714
DB 5411 AATATGGGATGATGTTAAGGCTATTGGTCTTTATGCTCTCTTGGTCGTCAGACTGATG 5470
QY 1715 GGCCCTATTCGGATATTGAGATGATGTGTCTGTCATGTCACAGAGGAAGCAGAGTTTCAGCC 1774
DB 5471 GGCCCTATTCGGATATTGAGATGATGTGTCTGTCATGTCACAGAGGAAGCAGAGTTTCAGCC 5530
QY 1775 ATGAATGGACAACCGGTGAGTGGAAGTGGAATTTTGTATGACGGAAGAGATTCTAC 1834
;

5531 ATGAATGGACAACCGGTGAGTGGAAGTGGAATTTTGTATGACGGAAGAGATTCTAC 5590
1835 TAGATTATGCATCTCAGGTGGAAATCAGATTGGCGGCTTACACATGGTCAATTTTCTCTA 1894
5591 TAGATTATGCATCTCAGGTGGAAATCAGATTGGCGGCTTACACATGGTCAATTTTCTCTA 5650
1895 TTTTGGCGAATTTATGATTCAAGTGGATACCTTAGAGAAAGTGATCAAACTGCTAAATCGG 1954
5651 TTTTGGCGAATTTATGATTCAAGTGGATACCTTAGAGAAAGTGATCAAACTGCTAAATCGG 5710
1955 TAGAAGCCCAACGTTCCACGATGCGATTGTGTCGCCCTTATCGTAGAAGAGCTGTTGAAT 2014
5711 TAGAAGCCCAACGTTCCACGATGCGATTGTGTCGCCCTTATCGTAGAAGAGCTGTTGAAT 5770
2015 ATGACAGCAAAATGCGGTAAATATTTCGTGTGCAAGCCGACCAACATTTCTACCATCCTTGA 2074
5771 ATGACAGCAAAATGCGGTAAATATTTCGTGTGCAAGCCGACCAACATTTCTACCATCCTTGA 5830
2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA 2134
5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA 5890
2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194
5891 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 5950
2195 ACCATCTGTGCCAGTTGCGTAATGTCGTGTCACACTTTCCGACTCTGAGAACTTCTCGAAT 2254
5951 ACCATCTGTGCCAGTTGCGTAATGTCGTGTCACACTTTCCGACTCTGAGAACTTCTCGAAT 6010
2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314
6011 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 6070
2315 ATGTGTCAAAAACGCATACCAATTTTGGAA 2341
6071 ATGTGTCAAAAACGCATACCAATTTTGGAA 6097

RESULT 9
US-08-470-369-21
; Sequence 21, Application US/08470369
; Patent No. 623889
; GENERAL INFORMATION:
; APPLICANT: Dorsers, Lambertus C. J.
; APPLICANT: Wagemaker, Gerard
; APPLICANT: Vos, Yvonne J.
; APPLICANT: Van Leen, Robert W.
; TITLE OF INVENTION: MOLECULAR CLONING AND EXPRESSION OF HUMAN
; TITLE OF INVENTION: IL-3
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,369
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,480
; FILING DATE:
; APPLICATION NUMBER: 07/854,297
; FILING DATE: 19-MAR-1992
; ATTORNEY/AGENT INFORMATION:
```

NAME: Gracey, Nancy J.
REGISTRATION NUMBER: 28,216
REFERENCE/DOCKET NUMBER: 24615-20002.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 7336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (plasmid)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: pBH1
US-08-470-369-21

Query Match 27.5%; Score 867; DB 3; Length 7336;
Best Local Similarity 100.0%; Pred. No. 2.6e-237; Indels 0; Gaps 0;
Matches 867; Conservative 0; Mismatches 0;

Qy 1475 TGCTTAGGAACAGAGTTATTAAATAGCTGAATAAGAACGGTCTCTCCAAATATTTCTAT 1534
Db |||||||
Qy 5231 TGCCTAGGAACAGAGTTATTAAATAGCTGAATAAGAACGGTCTCTCCAAATATTTCTAT 5290
Db |||||||
Qy 1535 TTAGAAAGCAATCTAAATATCTGAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1594
Db |||||||
Qy 5291 TTAGAAAGCAATCTAAATATCTGAAAGGGAATGAGAAATAGTGAATGGACCAATAA 5350
Db |||||||
Qy 1595 TAATGACTAGAGAAAGATGAGATGTTGTCATGAATTAAGAAAGCAATATGGATA 1654
Db |||||||
Qy 5351 TAATGACTAGAGAAAGATGAGATGTTGTCATGAATTAAGAAAGCAATATGGATA 5410
Db |||||||
Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATATGGCTCTCTGGTCTCAGACTGATG 1714
Db |||||||
Qy 5411 AATATGGGATGATGTTAAGGCTATTGGTGTATATGGCTCTCTGGTCTCAGACTGATG 5470
Db |||||||
Qy 1715 GGCCCTATTCGATATTGAGATGATGTTGTCATGTTCAACAGAGAAAGCAGATTACGCC 1774
Db |||||||
Qy 5471 GGCCCTATTCGATATTGAGATGATGTTGTCATGTTCAACAGAGAAAGCAGATTACGCC 5530
Db |||||||
Qy 1775 ATGATGACAAACCGGTGAGTGGAGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1834
Db |||||||
Qy 5531 ATGATGACAAACCGGTGAGTGGAGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 5590
Db |||||||
Qy 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
Db |||||||
Qy 5591 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 5650
Db |||||||
Qy 1895 TTTTGGCGATTTATGATTCAGGTGGATCTTATAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db |||||||
Qy 5651 TTTTGGCGATTTATGATTCAGGTGGATCTTATAGAGAAAGTGTATCAAACTGCTAAATCGG 5710
Db |||||||
Qy 1955 TAGAGCCCAACCGTTCACATCGGATTTGTCCTTATCGTAGAGAGCTGTTGAT 2014
Db |||||||
Qy 5711 TAGAGCCCAACCGTTCACATCGGATTTGTCCTTATCGTAGAGAGCTGTTGAT 5770
Db |||||||
Qy 2015 ATGAGGCAATGGCGTAATATTGTTGCAAGGACGACAAATTTCTACCATCCTTGA 2074
Db |||||||
Qy 5771 ATGAGGCAATGGCGTAATATTGTTGCAAGGACGACAAATTTCTACCATCCTTGA 5830
Db |||||||
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTGCATCATCGCATCTGTTATA 2134
Db |||||||
Qy 5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTGCATCATCGCATCTGTTATA 5890
Db |||||||
Qy 2135 CGAGCGGCTTCGCTCTTAATCAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db |||||||
Qy 5891 CGAGCGGCTTCGCTCTTAATCAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 5950
Db |||||||
Qy 2195 ACCATCTGTGCCAGTTGCTGTAATGTTGTTCAACTTTTCCGACTCTCGAGAAACTTCTGGAAT 2254
Db |||||||

Db 5951 ACCATCTGTGCCAGTTGCTGTAATGTTCTGCTCAACTTTCCGACTCTGAGAAACTTCTGGAAT 6010
Qy 2255 CGCTAGAGAATTTCTGGAATGGGATTCAGAGTGGAGACAGACAGCATATATAGTGG 2314
Db 6011 CGCTAGAGAATTTCTGGAATGGGATTCAGAGTGGAGACAGACAGCATATATAGTGG 6070
Qy 2315 ATGTGTCAAAACGCATACCATTTTGA 2341
Db 6071 ATGTGTCAAAACGCATACCATTTTGA 6097
RESULT 10
US-09-098-877B-3
; Sequence 3, Application US/09098877B
; Patent No. 6632633
; GENERAL INFORMATION:
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN, GERARDUS CORNELIS
; APPLICANT: MARIA; SMAAL, ERIC BASTIAN
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND
; TITLE OF INVENTION: GENETICALLY ENGINEERED CELLS USED THEREIN
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/098,877B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/418,085
; FILING DATE: 06-APR-1995
; APPLICATION NUMBER: US/08/054,185
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/002,608
; FILING DATE: 11-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/474,857
; FILING DATE: 30-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/474,798
; FILING DATE: 16-JULY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL89/00072
; FILING DATE: 25-SEPT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL/88/200904.6
; FILING DATE: 06-MAY-88
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL/88/202080.3
; FILING DATE: 03-SEP-88
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146.1169 CON-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 3:
; LENGTH: 7336
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE

; TOPOLOGY: UNKNOWN

; FEATURE:

; NAME/KEY:

; LOCATION:

; IDENTIFICATION METHOD:

; OTHER INFORMATION: PLASMID pBHA-1

US-09-098-877B-3

Query Match 27.5%; Score 867; DB 3; Length 7336;

Best Local Similarity 100.0%; Pred. No. 2.6e-237; Indels 0; Gaps 0;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGACAGATTATTAATAGCTGAATAGAAACGGTCTCTCAAATATTTCTAT 1534

Db 5231 TGCTTAGGAAGACAGATTATTAATAGCTGAATAGAAACGGTCTCTCAAATATTTCTAT 5290

Qy 1535 TTAGAAAGCAAACTCTAAATATCTGAAAGGAATGAGATAGTGAATGGAACCAATAA 1594

Db 5291 TTAGAAAGCAAACTCTAAATATCTGAAAGGAATGAGATAGTGAATGGAACCAATAA 5350

Qy 1595 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAACGAATATTGGATA 1654

Db 5351 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAACGAATATTGGATA 5410

Qy 1655 AATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTCGTCAGACTGATG 1714

Db 5411 AATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTCGTCAGACTGATG 5470

Qy 1715 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCATGTCACAGAGGAACGAGTTTCAGCC 1774

Db 5471 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCATGTCACAGAGGAACGAGTTTCAGCC 5530

Qy 1775 ATGAATGCACAAACCGGTGAGTGAAGGTGGAAGTGAATTTTGTAGCGAAGAGATTTCTAC 1834

Db 5531 ATGAATGCACAAACCGGTGAGTGAAGGTGGAAGTGAATTTTGTAGCGAAGAGATTTCTAC 5590

Qy 1835 TAGATTAATGATCTCAGGTGGAATCAGATTGGCCGCTTACATGTCGTCATTTTCTCTA 1894

Db 5591 TAGATTAATGATCTCAGGTGGAATCAGATTGGCCGCTTACATGTCGTCATTTTCTCTA 5650

Qy 1895 TTTTGCCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954

Db 5651 TTTTGCCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 5710

Qy 1955 TAGAAGCCAAACCGTTCACGATCCGATTTGTGCGCTTATCGTAGAAGAGCTGTTGCAAT 2014

Db 5711 TAGAAGCCAAACCGTTCACGATCCGATTTGTGCGCTTATCGTAGAAGAGCTGTTGCAAT 5770

Qy 2015 ATGACGGCAAAATGCGGTAATAATTCGTGTGCAAGACCGACAAACATTTCTACCATCTGA 2074

Db 5771 ATGACGGCAAAATGCGGTAATAATTCGTGTGCAAGACCGACAAACATTTCTACCATCTGA 5830

Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2134

Db 5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 5890

Qy 2135 CGACGACGCTTCGGTCTTAACTCAACGAGTTAAGCAATCAGATCTTCTTCAGTTATG 2194

Db 5891 CGACGACGCTTCGGTCTTAACTCAACGAGTTAAGCAATCAGATCTTCTTCAGTTATG 5950

Qy 2195 ACCATCTGTGCGAGTTCGTAAATGTCGTGCTCAACTTTTCCGACTCTGAGAACTTCGGAAT 2254

Db 5951 ACCATCTGTGCGAGTTCGTAAATGTCGTGCTCAACTTTTCCGACTCTGAGAACTTCGGAAT 6010

Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314

Db 6011 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 6070

Qy 2315 ATGTGTCAAAACGCATACCATTTTGA 2341

Db 6071 ATGTGTCAAAACGCATACCATTTTGA 6097

RESULT 11

US-08-460-343B-1

; Sequence 1, Application US/08460343B

; Patent No. 5741664

; GENERAL INFORMATION:

; APPLICANT: Marcus D. Ballinger and James A. Wells

; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING

; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING

; NUMBER OF SEQUENCES: 74

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 1 DNA Way

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/460,343B

; FILING DATE: 01-Jun-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/398028

; FILING DATE: 03-mar-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Kubinec, Jeffrey S.

; REGISTRATION NUMBER: 36,575

; REFERENCE/DOCKET NUMBER: P0936C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-8228

; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8119 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

US-08-460-343B-1

Query Match 27.5%; Score 867; DB 2; Length 8119;

Best Local Similarity 100.0%; Pred. No. 2.8e-237;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGACAGATTATTAATAGCTGAATAGAAACGGTCTCTCAAATATTTCTAT 1534

Db 5764 TGCTTAGGAAGACAGATTATTAATAGCTGAATAGAAACGGTCTCTCAAATATTTCTAT 5823

Qy 1535 TTAGAAAGCAAACTCTAAATATCTGAAAGGAATGAGATAGTGAATGGAACCAATAA 1594

Db 5824 TTAGAAAGCAAACTCTAAATATCTGAAAGGAATGAGATAGTGAATGGAACCAATAA 5883

Qy 1595 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAACGAATATTGGATA 1654

Db 5884 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAACGAATATTGGATA 5943

Qy 1655 AATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTCGTCAGACTGATG 1714

Db 5944 AATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTCGTCAGACTGATG 6003

Qy 1715 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCATGTCACAGAGGAACGAGTTTCAGCC 1774

Db 6004 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAGGAACGAGTTTCAGCC 6063

Qy 1775 ATGAATGGACAAACCGGTGAGTGAAGGTGGAAGTGAATTTTGTAGCGAAGAGATTTCTAC 1834

Db 6064 ATGAATGGACAAACCGGTGAGTGAAGGTGGAAGTGAATTTTGTAGCGAAGAGATTTCTAC 6123

Qy 1835 TAGATTAATGATCTCAGGTGGAATCAGATTGGCGCTTACATGTTGTTTCTCTA 1894

Db 6071 TAGATTAATGATCTCAGGTGGAATCAGATTGGCGCTTACATGTTGTTTCTCTA 1894

Db 6124 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTTACACATGGTCAATTTTCTCTA 6183
Qy 1895 TTTTGGCGATTTATGATTCAGGTGGATCTTTAGAGAAAGTGATCAAACTCTCTAAATCGG 1954
Db 6184 TTTTGGCGATTTATGATTCAGGTGGATCTTTAGAGAAAGTGATCAAACTCTCTAAATCGG 6243
Qy 1955 TAGAAGCCCAACCGTTCCACGATGCGATTTCTGCGCTTATCGTAGAAGAGCTGTTTGAAT 2014
Db 6244 TAGAAGCCCAACCGTTCCACGATGCGATTTCTGCGCTTATCGTAGAAGAGCTGTTTGAAT 6303
Qy 2015 ATGCAGGCAAAATGGCGTAAATTTCTGTGTCAGAGCCGACCAACATTTCTACCATCCTTGA 2074
Db 6304 ATGCAGGCAAAATGGCGTAAATTTCTGTGTCAGAGCCGACCAACATTTCTACCATCCTTGA 6363
Qy 2075 CTGTACAGGTAGCAATGCGAGTGGCATTTGATTTGATTTGCTGTCATCGCATCTGTTATA 2134
Db 6364 CTGTACAGGTAGCAATGCGAGTGGCATTTGATTTGATTTGCTGTCATCGCATCTGTTATA 6423
Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAAGCAATCAGATCTTCTTTCAGGTTATG 2194
Db 6424 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAAGCAATCAGATCTTCTTTCAGGTTATG 6483
Qy 2195 ACCATCTGTGCCAGTTCTGATATGCTCGTCAACTTTCCGACTCTGAGAAACTTCTGGAAT 2254
Db 6484 ACCATCTGTGCCAGTTCTGATATGCTCGTCAACTTTCCGACTCTGAGAAACTTCTGGAAT 6543
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACCGATATATAGTGG 2314
Db 6544 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACCGATATATAGTGG 6603
Qy 2315 ATGTGTCAAAACGCATACCATTTTGA 2341
Db 6604 ATGTGTCAAAACGCATACCATTTTGA 6630

RESULT 12

US-08-398-028B-1

; Sequence 1, Application US/08398028B

; Patent No. 5780285

; GENERAL INFORMATION:

; APPLICANT: Marcus D. Ballinger and James A. Wells

; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING

; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES

; NUMBER OF SEQUENCES: 74

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 1 DNA Way

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA: US/08/398,028B

; FILING DATE: 03-Mar-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Kubinec, Jeffrey S.

; REGISTRATION NUMBER: 36,575

; REFERENCE/DOCKET NUMBER: P0936

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-8228

; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8119 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

US-08-398-028B-1

Query Match 27.5%; Score 867; DB 2; Length 8119;

Best Local Similarity 100.0%; Pred. No. 2.8e-237;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGTCTCTCAAAATATTTCTTAT 1534

Db 5764 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGTCTCTCAAAATATTTCTTAT 5823

Qy 1535 TTAGAAAAGCAAACTCTAAAATTTATCTCAAAAGGAATGAGAAATAGTGAATGGACCAATAA 1594

Db 5824 TTAGAAAAGCAAACTCTAAAATTTATCTCAAAAGGAATGAGAAATAGTGAATGGACCAATAA 5883

Qy 1595 TAATGACTAGAGAAAGAAATGAAGATTTCTCATGAAATTAAGGAACGATATTTGGATA 1654

Db 5884 TAATGACTAGAGAAAGAAATGAAGATTTCTCATGAAATTAAGGAACGATATTTGGATA 5943

Qy 1655 AATATGGGATGATTTAAGGCTATTGGTGTCTTATGGCTCTCTTGGTCTGTCAGACTGATG 1714

Db 5944 AATATGGGATGATTTAAGGCTATTGGTGTCTTATGGCTCTCTTGGTCTGTCAGACTGATG 6003

Qy 1715 GGCCCTATTCCGATTTGAGATGATGTGTCTCATGTCAACAGAGAAAGAGATTCAGCC 1774

Db 6004 GGCCCTATTCCGATTTGAGATGATGTGTCTCATGTCAACAGAGAAAGAGATTCAGCC 6063

Qy 1775 ATGAATGGACAAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGCGAGAGATTCCTAC 1834

Db 6064 ATGAATGGACAAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGCGAGAGATTCCTAC 6123

Qy 1835 TAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTTACACATGGTCAATTTTCTCTA 1894

Db 6124 TAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTTACACATGGTCAATTTTCTCTA 6183

Qy 1895 TTTTGGCGATTTATGATTCAGGTGGATPACTTTAGAGAAAGTGTATCAAACTCTAAATCGG 1954

Db 6184 TTTTGGCGATTTATGATTCAGGTGGATPACTTTAGAGAAAGTGTATCAAACTCTAAATCGG 6243

Qy 1955 TAGAAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014

Db 6244 TAGAAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 6303

Qy 2015 ATGCAGGCAAAATGGCGTAAATTTCTGTGCAAGGACCGACCAACATTTCTACCATCCTTGA 2074

Db 6304 ATGCAGGCAAAATGGCGTAAATTTCTGTGCAAGGACCGACCAACATTTCTACCATCCTTGA 6363

Qy 2075 CTGTACAGGTAGCAATGGCAGTGGCCATGTTGATTTGATTTGCTGTCATCGCATCTGTTATA 2134

Db 6364 CTGTACAGGTAGCAATGGCAGTGGCCATGTTGATTTGATTTGCTGTCATCGCATCTGTTATA 6423

Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTTCAGGTTATG 2194

Db 6424 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTTCAGGTTATG 6483

Qy 2195 ACCATCTGTGCCAGTTCTGATATGCTCGTCAACTTTCCGACTCTGAGAAACTTCTGGAAT 2254

Db 6484 ACCATCTGTGCCAGTTCTGATATGCTCGTCAACTTTCCGACTCTGAGAAACTTCTGGAAT 6543

Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACCGATATATAGTGG 2314

Db 6544 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACCGATATATAGTGG 6603

RESULT 13

US-08-504-265B-1

; Sequence 1, Application US/08504265B

; Patent No. 5837516

; GENERAL INFORMATION:

; APPLICANT: Marcus D. Ballinger and James A. Wells

;; TITLE OF INVENTION: SUBTILISIN VARIANTS CAPABLE OF CLEAVING
;; TITLE OF INVENTION: SUBSTRATES CONTAINING BASIC RESIDUES
;; NUMBER OF SEQUENCES: 90
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 1 DNA Way
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WinPatIn (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/504,265B
;; FILING DATE: 19-Jul-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/398028
;; FILING DATE: 03-MAR-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kubinec, Jeffrey S.
;; REGISTRATION NUMBER: 36,575
;; REFERENCE/DOCKET NUMBER: P0936P1
;; TELEPHONE: 650/952-8228
;; TELEFAX: 650/952-9881
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 819 base pairs
;; TYPE: Nucleic Acid
;; STRANDEDNESS: Single
;; TOPOLOGY: Linear
;; US-08-504-265B-1

Query Match 27.5%; Score 867; DB 2; Length 8119;
Best Local Similarity 100.0%; Pred. No. 2.8e-237;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1475 TGCTTAGAAGACAGGTTATTAATAGCTGAATGAAGACGGTCTCTCCAAATATTTCTAT 1534
DB 5764 TGCTTAGAAGACAGGTTATTAATAGCTGAATGAAGACGGTCTCTCCAAATATTTCTAT 5823
QY 1535 TTAGAAAGCAATCTAAATTTATCTGAAAGGAATGAAGTAGTGAATGGACCAATAA 1594
DB 5824 TTAGAAAGCAATCTAAATTTATCTGAAAGGAATGAAGTAGTGAATGGACCAATAA 5883
QY 1595 TAATGACTAGAGAAGAAAGAAATGAAGATTCTTCATGAAATTAAGGAACGAATATTGGATA 1654
DB 5884 TAATGACTAGAGAAGAAAGAAATGAAGATTCTTCATGAAATTAAGGAACGAATATTGGATA 5943
QY 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATATGGCTCTCTTGGTCTGACAGTGAATG 1714
DB 5944 AATATGGGATGATGTTAAGGCTATTGGTGTATATGGCTCTCTTGGTCTGACAGTGAATG 6003
QY 1715 GGCCCTATTCCGATATTGAGATGATGTCATGTCACAGAGGAGAGAGTTCAGCC 1774
DB 6004 GGCCCTATTCCGATATTGAGATGATGTCATGTCACAGAGGAGAGAGTTCAGCC 6063
QY 1775 ATGAATGGACAACCGGTGAGTGAAGGTGAAGTGAATTTGATAGCGAAGAGATTCTAC 1834
DB 6064 ATGAATGGACAACCGGTGAGTGAAGGTGAAGTGAATTTGATAGCGAAGAGATTCTAC 6123
QY 1835 TAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACATGAGTCAATTTTCTCTA 1894
DB 6124 TAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACATGAGTCAATTTTCTCTA 6183
QY 1895 TTTTGGCGATTATGATTCAGGTGCGATACATTAGGAAGGTATCAAACTGCTAAATCGG 1954
DB 6184 TTTTGGCGATTATGATTCAGGTGCGATACATTAGGAAGGTATCAAACTGCTAAATCGG 6243

QY 1955 TAGAAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014
DB 6244 TAGAAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 6303
QY 2015 ATGACAGGCAAAATGGCGTAAATATTCGTTGTCGAAGGACCGCAACATTTCTACCATCCTTGA 2074
DB 6304 ATGACAGGCAAAATGGCGTAAATATTCGTTGTCGAAGGACCGCAACATTTCTACCATCCTTGA 6363
QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGTCATCATCGCATCTGTTTATA 2134
DB 6364 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGTCATCATCGCATCTGTTTATA 6423
QY 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCCTCAGGTATG 2194
DB 6424 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCCTCAGGTATG 6483
QY 2195 ACCATCTGTGCCAGTTCCGTAATCTCTGGTCAACTTTCCGACTCTGAGAACTTCTCGAAT 2254
DB 6484 ACCATCTGTGCCAGTTCCGTAATCTCTGGTCAACTTTCCGACTCTGAGAACTTCTCGAAT 6543
QY 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGAGATATATAGTGG 2314
DB 6544 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGAGATATATAGTGG 6603
QY 2315 ATGTGTCAAAACGCATACCATTTTGA 2341
DB 6604 ATGTGTCAAAACGCATACCATTTTGA 6630

RESULT 14
US-08-556-978B-79/c
; Sequence 79, Application US/08556978B
; Patent No. 6268169
; GENERAL INFORMATION:
; APPLICANT: FAHNESTOCK, STEPHEN F.
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,978B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,600
; FILING DATE: JUNE 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9389-A
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-556-978B-79

Search completed: February 11, 2006, 23:26:19
Job time : 403 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2006, 02:30:24 ; Search time 1626 Seconds
(without alignments)
16050.527 Million cell updates/sec

Title: US-10-681-086-1
Perfect score: 3156
Sequence: 1 ggatccacgaggttacgagc.....atcggtatgatgcgaattc 3156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3156	100.0	3156	7	US-10-681-086-1
2	1219	38.6	3560	9	US-10-505-855-1
3	1081	34.3	1170	6	US-10-369-493-46942
4	867	27.5	5793	3	US-09-869-855A-3
5	867	27.5	5865	7	US-10-402-466A-2
6	867	27.5	5943	3	US-09-869-855A-2
7	867	27.5	6694	7	US-10-402-466A-1
8	867	27.5	7336	5	US-10-084-814-3
9	867	27.5	7336	7	US-10-402-013-4
10	867	27.5	7336	7	US-10-462-128-3
11	867	27.5	8198	7	US-10-402-466A-7
12	867	27.5	9286	7	US-10-402-466A-10
13	867	27.5	10929	5	US-10-032-393-7
14	865.4	27.4	5302	7	US-10-385-415-2
15	865.4	27.4	5767	7	US-10-385-415-4
16	788.8	24.4	1365	8	US-10-479-674-229
17	759	22.0	759	7	US-10-718-628-10
18	701	22.2	6837	3	US-09-928-847B-49
19	521.2	16.5	1133	3	US-09-974-300-993
20	494.4	15.7	992	3	US-09-974-300-980
21	471	14.9	6405	3	US-09-869-855A-1
22	285.6	9.0	996	7	US-10-282-122A-9649
23	285	9.0	963	7	US-10-282-122A-35025

ALIGNMENTS

RESULT 1
US-10-681-086-1
; Sequence 1, Application US/10681086
; Publication No. US20040063171A1
; GENERAL INFORMATION:
; APPLICANT: HOHMANN Dr., Hans-Peter
; APPLICANT: MOUNCEY Dr., Nigel J.
; APPLICANT: SCHLEIKER Dr., Heinrich W.
; APPLICANT: STEBBINS Dr., Jeffrey W.
; TITLE OF INVENTION: PROCESS FOR PRODUCING A TARGET PERMENTATION PRODUCT
; FILE REFERENCE: Process For Prod. Target Ferm. Product
; CURRENT APPLICATION NUMBER: US/10/681,086
; CURRENT FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PRIOR APPLICATION NUMBER: US/09/633,927
; PRIOR FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3156
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-10-681-086-1
Query Match 100.0%; Score 3156; DB 7; Length 3156;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGA 60
Db 1 GGATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGA 60
Qy 61 ATGAAGAAGCCGCGGTACATCGTAACCTCGGTCAATGATGCGCGCCGTTCCAGAG 120
Db 61 ATGAAGAAGCCGCGGTACATCGTAACCTCGGTCAATGATGCGCGCCGTTCCAGAG 120
Qy 121 AGGAATATTGATGCGGAAATCAAAACGGTCTGGTCTCTCAACCAATTTATTTAGGGCTCGCA 180
Db 121 AGGAATATTGATGCGGAAATCAAAACGGTCTGGTCTCTCAACCAATTTATTTAGGGCTCGCA 180
Qy 181 AGCGATAGAGTTTGATTCGATCGAGCCCAACAGCATTCGAGCAATTTGGGACAGGAAGC 240
Db 181 AGCGATAGAGTTTGATTCGATCGAGCCCAACAGCATTCGAGCAATTTGGGACAGGAAGC 240
Qy 241 AGCGTTTCAAGTTTAAACGACAGCAATTCGGTCTGGCATGAAAAGCTAGAAAAGAGATT 300
Db 241 AGCGTTTCAAGTTTAAACGACAGCAATTCGGTCTGGCATGAAAAGCTAGAAAAGAGATT 300

QY 301 GCCAGCTTTAAACTGACAGAAACGCGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTC 360
DB 301 GCCAGCTTTAACTGACAGAAACGCGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTC 360
QY 361 GGTGTCCTTTTCATCCTTCCGAGAAAGGAAGATGTCAFTTTTAAAGTGACAGCTCAATCAT 420
DB 361 GGTGTCCTTTTCATCCTTCCGAGAAAGGAAGATGTCAFTTTTAAAGTGACAGCTCAATCAT 420
QY 421 GCAAGTAGATCGAGCGCTGCGGACTTTCCTAAGGCTGTATACAGTTGTTTATTCGGCATATT 480
DB 421 GCAAGTAGATCGAGCGCTGCGGACTTTCCTAAGGCTGTATACAGTTGTTTATTCGGCATATT 480
QY 481 GATATGAATGATCTTGAATAACAGCTGAATGAAACACAGCGGTATACAGCGCGGTTTATC 540
DB 481 GATATGAATGATCTTGAATAACAGCTGAATGAAACACAGCGGTATACAGCGCGGTTTATC 540
QY 541 GTAACAGACGGAGTATTCAGCATGGAATGACACAAATCGCCCTCTTTGATCAGATCATCTCA 600
DB 541 GTAACAGACGGAGTATTCAGCATGGAATGACACAAATCGCCCTCTTTGATCAGATCATCTCA 600
QY 601 CTTGGAAACGCTATCATGCTTGGTGTGCTGATGATGATGCCACGCAACAGGAGTTTTCG 660
DB 601 CTTGGAAACGCTATCATGCTTGGTGTGCTGATGATGATGCCACGCAACAGGAGTTTTCG 660
QY 661 GCGGATTCGGGACAAAGGAACGAGTGAATACCTTTTGGTGTGCTGCGCATTTGTTATCGGC 720
DB 661 GCGGATTCGGGACAAAGGAACGAGTGAATACCTTTTGGTGTGCTGCGCATTTGTTATCGGC 720
QY 721 ACCTTAAAGCAAGCTGTTGGCGCGGAAGGAGGTTTTCGGCAGGATCAGCGGCTCTTCATC 780
DB 721 ACCTTAAAGCAAGCTGTTGGCGCGGAAGGAGGTTTTCGGCAGGATCAGCGGCTCTTCATC 780
QY 781 GACTTTTGTGTAACCATGCCAGAACATTTATCTTTTAAACGGCTATTCGCCACGCGCAGC 840
DB 781 GACTTTTGTGTAACCATGCCAGAACATTTATCTTTTAAACGGCTATTCGCCACGCGCAGC 840
QY 841 TGTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGCGGAAACAGCAGCTT 900
DB 841 TGTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGCGGAAACAGCAGCTT 900
QY 901 TTATTTTCTTATATCAGCATGATCAGAACCAAGTCTGAAGAAATATGGGTATGTGGTGA 960
DB 901 TTATTTTCTTATATCAGCATGATCAGAACCAAGTCTGAAGAAATATGGGTATGTGGTGA 960
QY 961 GGAGATCACACCGATTTATTCCTGTAGTCAATGGCGATGCCATATAAACCGGCTCTATT 1020
DB 961 GGAGATCACACCGATTTATTCCTGTAGTCAATGGCGATGCCATATAAACCGGCTCTATT 1020
QY 1021 GCTGAAAACCTGCAGGCAAGGAATTTATGCTCTGCTCATTTGCGCGCCCAACCGTTGCG 1080
DB 1021 GCTGAAAACCTGCAGGCAAGGAATTTATGCTCTGCTCATTTGCGCGCCCAACCGTTGCG 1080
QY 1081 CCGGCTGAAAGCCGGATTCGAAGCTTGGGACAGAGGTGCGATCAGGGAATGAGTTTATA 1140
DB 1081 CCGGCTGAAAGCCGGATTCGAAGCTTGGGACAGAGGTGCGATCAGGGAATGAGTTTATA 1140
QY 1141 AAATAAAAAAGCACTGAAAAAGGTGCTTTTTTTTGGTGGTTTGAACCTGTTCTTTCTT 1200
DB 1141 AAATAAAAAAGCACTGAAAAAGGTGCTTTTTTTTGGTGGTTTGAACCTGTTCTTTCTT 1200
QY 1201 ATCTTTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTAGTTGCTGAAAGGTGCG 1260
DB 1201 ATCTTTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTAGTTGCTGAAAGGTGCG 1260
QY 1261 TTGAAGTGTGGTATGATGTTTTTAAAGTATTTGAAACCCCTTAAAAATGGTTGCGACAG 1320
DB 1261 TTGAAGTGTGGTATGATGTTTTTAAAGTATTTGAAACCCCTTAAAAATGGTTGCGACAG 1320
QY 1321 AAAAAACCCCATCTGTTAAAGTTATAAGTGACTATAACAAATAAATAAGATGGGGTTTT 1380
DB 1321 AAAAAACCCCATCTGTTAAAGTTATAAGTGACTATAACAAATAAATAAGATGGGGTTTT 1380

QY 1381 CTTTAAATATATTGTCCTTAATAGTAGCATTTTATTTCAGATGAAAAATCAAGGGTTTTAG 1440
DB 1381 CTTTAAATATATTGTCCTTAATAGTAGCATTTTATTTCAGATGAAAAATCAAGGGTTTTAG 1440
QY 1441 TGGACAAGACAAAAGTGGAAAAAGTGAGACCAATGTCTTAGGAAAGACGAGTTATTAATAG 1500
DB 1441 TGGACAAGACAAAAGTGGAAAAAGTGAGACCAATGTCTTAGGAAAGACGAGTTATTAATAG 1500
QY 1501 CTGNAATGAAGACGGTCTCTCCAAATATTCTTATTTAGAAAAGCAAAATCTAAAAATATCT 1560
DB 1501 CTGNAATGAAGACGGTCTCTCCAAATATTCTTATTTAGAAAAGCAAAATCTAAAAATATCT 1560
QY 1561 GAAAAAGGAATGAGAAATAGTGAATGGACCAATATAATAGTACAGAGAAAGAAATGAAG 1620
DB 1561 GAAAAAGGAATGAGAAATAGTGAATGGACCAATATAATAGTACAGAGAAAGAAATGAAG 1620
QY 1621 ATTGTTTCATGAAATTAAGGAACGAATATTGGATAAATATGGGGATGATGTTTAAAGGCTATT 1680
DB 1621 ATTGTTTCATGAAATTAAGGAACGAATATTGGATAAATATGGGGATGATGTTTAAAGGCTATT 1680
QY 1681 GGTGTTTTATGGCTCTCTTGGTCTGACACTGATGGGCCCTATTTCGGATATTTCAGATGATG 1740
DB 1681 GGTGTTTTATGGCTCTCTTGGTCTGACACTGATGGGCCCTATTTCGGATATTTCAGATGATG 1740
QY 1741 TGTGTCATGTCACACAGAGGAAGCAGAGTTTCAAGCCATGAATGGACAAACCGGTGAGTGGAAAG 1800
DB 1741 TGTGTCATGTCACACAGAGGAAGCAGAGTTTCAAGCCATGAATGGACAAACCGGTGAGTGGAAAG 1800
QY 1801 GTGGAAGTGAATTTTGTATAGCGAAGAGATTTCTACTAGATTATGCAATCTCAGGTTGGAATCA 1860
DB 1801 GTGGAAGTGAATTTTGTATAGCGAAGAGATTTCTACTAGATTATGCAATCTCAGGTTGGAATCA 1860
QY 1861 GATTGGCCGCTTACACATGCTCAATTTTCTTATTTTGGATTTTATGATTCAGGTGGA 1920
DB 1861 GATTGGCCGCTTACACATGCTCAATTTTCTTATTTTGGCATTTTATGATTCAGGTGGA 1920
QY 1921 TACTTAGAAGAAAGTGTATCAAACTGCTAAATCGGTAGAAGCCCAAAAGCTTCCACGATGCG 1980
DB 1921 TACTTAGAAGAAAGTGTATCAAACTGCTAAATCGGTAGAAGCCCAAAAGCTTCCACGATGCG 1980
QY 1981 ATTTGTCCTTATCGTAGAAGAGCTGTTTGAATATATCAGGCAAAATGGCGTAAATTTCTGT 2040
DB 1981 ATTTGTCCTTATCGTAGAAGAGCTGTTTGAATATATCAGGCAAAATGGCGTAAATTTCTGT 2040
QY 2041 GTGCAAGGACCGCAACATTTCTACCATCTGCTGATCAGAGTAGCAATGGCAGGTGCC 2100
DB 2041 GTGCAAGGACCGCAACATTTCTACCATCTGCTGATCAGAGTAGCAATGGCAGGTGCC 2100
QY 2101 ATGTTGATTGGTCTGCAATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
DB 2101 ATGTTGATTGGTCTGCAATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
QY 2161 GCAGTTAAGCAATCAGATCTTCTTCAAGTTATGACCAATCTGTGCCAGTTCTGTAATGTCT 2220
DB 2161 GCAGTTAAGCAATCAGATCTTCTTCAAGTTATGACCAATCTGTGCCAGTTCTGTAATGTCT 2220
QY 2221 GGTCAACCTTTCCGACTCTGAGAACTTCTCGAATCGCTAGAGATTTTCTGGAATGGGAT 2280
DB 2221 GGTCAACCTTTCCGACTCTGAGAACTTCTCGAATCGCTAGAGATTTTCTGGAATGGGAT 2280
QY 2281 CAGGAGTGGACAGAAACGACCGGATATATAGTGGATGTGTCAAAACGCATACCATTTTGA 2340
DB 2281 CAGGAGTGGACAGAAACGACCGGATATATAGTGGATGTGTCAAAACGCATACCATTTTGA 2340
QY 2341 ATTGCAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGGCG 2400
DB 2341 ATTGCAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGGCG 2400
QY 2401 CGAAGCGGCGCAGATCTGAATATTCGGCACATATTCTATCGTCGCAAGCGGACAGAGTCT 2460
DB 2401 CGAAGCGGCGCAGATCTGAATATTCGGCACATATTCTATCGTCGCAAGCGGACAGAGTCT 2460
QY 2461 CGTCTAACAGAGAAGTGGATCGGTCGTAGATCGGTTTCAGGAAATTTAAAGAGACGTATG 2520

Db 2461 CGTCTAACAGAGAAGTGGATCAGGTCTAGATCGGTTTCAGGAAATTAAGAGACGCTATG 2520
Qy 2521 GACTGAGATTTGTGCATGCTTTGGACTGTTGAAGCCAGAGCAGGCGAAGCGCTCAAAG 2580
Db 2521 GACTGAGATTTGTGCATGCTTTGGACTGTTGAAGCCAGAGCAGGCGAAGCGCTCAAAG 2580
Qy 2581 ATGACGAGTAGACCGCTATAATCATATTTGAATACGTACAGAGAAACCATTCRAACA 2640
Db 2581 ATGACGAGTAGACCGCTATAATCATATTTGAATACGTACAGAGAAACCATTCRAACA 2640
Qy 2641 TCACAACCTCACATACATACATACATACATACATACATACATACATACATACATACAT 2700
Db 2641 TCACAACCTCACATACATACATACATACATACATACATACATACATACATACATACAT 2700
Qy 2701 GGCTGCTCCGTGTTCAAGCGCCATATTCGGATGAGGAGACGAAACAGGATGTCATTG 2760
Db 2701 GGCTGCTCCGTGTTCAAGCGCCATATTCGGATGAGGAGACGAAACAGGATGTCATTG 2760
Qy 2761 ACATCGCAAAAGCTTGAAGGCTCTTGACGCGGATTCATTCCTGTGAAATTTTGTGATG 2820
Db 2761 ACATCGCAAAAGCTTGAAGGCTCTTGACGCGGATTCATTCCTGTGAAATTTTGTGATG 2820
Qy 2821 CAATTGATGGACCGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 2880
Db 2821 CAATTGATGGACCGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 2880
Qy 2881 TGCTGGCGCTGTTCCGCTTTATCAATCCATCAAGGAAATTCGATTTCCGGAGGAGAG 2940
Db 2881 TGCTGGCGCTGTTCCGCTTTATCAATCCATCAAGGAAATTCGATTTCCGGAGGAGAG 2940
Qy 2941 AGGTCAATCTCCGACATTCGAGCCTATAGGCTTTACGGCGCAAACTCCATTTTGTGCG 3000
Db 2941 AGGTCAATCTCCGACATTCGAGCCTATAGGCTTTACGGCGCAAACTCCATTTTGTGCG 3000
Qy 3001 GAGACTACTTAAACAACTGCCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Db 3001 GAGACTACTTAAACAACTGCCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Qy 3061 TAGGCTTTGAAGTTGAATCAGTGAAGAAATGAAGGCTAGTTAAAGTGCAGAAAGCTGAA 3120
Db 3061 TAGGCTTTGAAGTTGAATCAGTGAAGAAATGAAGGCTAGTTAAAGTGCAGAAAGCTGAA 3120
Qy 3121 AGAATCAATAAAGCAATCGTATGATGTCGAATTC 3156
Db 3121 AGAATCAATAAAGCAATCGTATGATGTCGAATTC 3156

RESULT 2

US-10-505-855-1
; Sequence 1, Application US/10505855
; Publication No. US20050124030A1
; GENERAL INFORMATION:
; APPLICANT: Roche Vitamins AG
; TITLE OF INVENTION: Fermentation Process
; FILE REFERENCE: 21154
; CURRENT APPLICATION NUMBER: US/10/505,855
; CURRENT FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3560
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(3560)
; OTHER INFORMATION: plasmid pNMR21
US-10-505-855-1

Query Match 38.6%; Score 1219; DB 9; Length 3560;
Best Local Similarity 99.5%; Pred. No. 1.3e-295;
Matches 1235; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1102 AGCTTGGCGACGAGGTCCGAGATCAGGGAATGAGTTTATAAAATAAAAAAGCACCTGAAA 1161
Db 1466 AGCTTGGCGACGAGGTCCGAGATCAGGGAATGAGTTTATAAAATAAAAAAGCACCTGAAA 1525
Qy 1162 AGGTGCTCTTTTGTGATGTTTGAACCTGTTCTTCTTATCTTGTATCATATGATATAGAATA 1221
Db 1526 AGGTGCTCTTTTGTGATGTTTGAACCTGTTCTTCTTATCTTGTATCATATGATATAGAATA 1585
Qy 1222 ACCTCATTTTATATTTTTATTTAGTTGCTGAAAGGTGCGTTGAAGTGTGCTGATGATGT 1281
Db 1586 ACCTC-----ATTATTTTATTTAGTTGCTGAAAGGTGCGTTGAAGTGTGCTGATGATGT 1639
Qy 1282 GTTTTAAAGTATTGAAAAACCTTTAAATTTGGTTTGCACAGAAAAACCCCATCTGTGTTAAAGT 1341
Db 1640 GTTTTAAAGTATTGAAAAACCTTTAAATTTGGTTTGCACAGAAAAACCCCATCTGTGTTAAAGT 1699
Qy 1342 TATAAGTACTAAACAAATAAATAAGATAGAGGGGTTTCTTTAATAATTTATGTTGCTCTA 1401
Db 1700 TATAAGTACTAAACAAATAAATAAGATAGAGGGGTTTCTTTAATAATTTATGTTGCTCTA 1759
Qy 1402 ATAGTAGCATTTATTCAGATGAAAAATCAAGGGTTTGTAGTGGCAAGACAAAAAGTCGAA 1461
Db 1760 ATAGTAGCATTTATTCAGATGAAAAATCAAGGGTTTGTAGTGGCAAGACAAAAAGTCGAA 1819
Qy 1462 AAGTGAGACCATGCTGCTTAGGAAAGACGAGTTATTAATAGCTGAATAAGAACGGTGTCTC 1521
Db 1820 AAGTGAGACCATGCTGCTTAGGAAAGACGAGTTATTAATAGCTGAATAAGAACGGTGTCTC 1879
Qy 1522 CAAATATCTTATTTAGAAAGCAAAATCTAAATTTATCTGAAAGGGAAATGAGNAATAGTG 1581
Db 1880 CAAATATCTTATTTAGAAAGCAAAATCTAAATTTATCTGAAAGGGAAATGAGNAATAGTG 1939
Qy 1582 AATGGACCAATATATGACTAGAGAGAAAGAAATGAAGATTTCTCATGAAATTTAAAGAA 1641
Db 1940 AATGGACCAATATATGACTAGAGAGAAAGAAATGAAGATTTCTCATGAAATTTAAAGAA 1999
Qy 1642 CGAATATTTGATAAATATGCGGATGATTTAAGGCTATTTGGTGTGTTATGGCTCTCTTGGT 1701
Db 2000 CGAATATTTGATAAATATGCGGATGATTTAAGGCTATTTGGTGTGTTATGGCTCTCTTGGT 2059
Qy 1702 CGTCAGATGATGGGCGCTTATCGGATATGAGATGATGTTGTCATGTCACACAGAGAA 1761
Db 2060 CGTCAGATGATGGGCGCTTATCGGATATGAGATGATGTTGTCATGTCACACAGAGAA 2119
Qy 1762 GCAGAGTTCAGCCATGAATGCGCAACCGGTCAGTGGAAAGTGAAGTATTTTGTATAGC 1821
Db 2120 GCAGAGTTCAGCCATGAATGCGCAACCGGTCAGTGGAAAGTGAAGTATTTTGTATAGC 2179
Qy 1822 GAAGAGATTTCTACTAGATTTATGATCTCAGGTGGAATCAGATTTGGCGCTTTACACATGGT 1881
Db 2180 GAAGAGATTTCTACTAGATTTATGATCTCAGGTGGAATCAGATTTGGCGCTTTACACATGGT 2239
Qy 1882 CAAATTTTCTTATTTTGGCGATTTATGATCAGGTGGAATCAGTGTAGAGAAAGTGTATCAA 1941
Db 2240 CAAATTTTCTTATTTTGGCGATTTATGATCAGGTGGAATCAGTGTAGAGAAAGTGTATCAA 2299
Qy 1942 ACTGCTAAATCGGTAGAACCCCAAGCTTCCAGATCGGATTTGTGCCCTTATCGTAGAA 2001
Db 2300 ACTGCTAAATCGGTAGAACCCCAAGCTTCCAGATCGGATTTGTGCCCTTATCGTAGAA 2359
Qy 2002 GAGCTGTTTGAATATGAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACAACTTT 2061
Db 2360 GAGCTGTTTGAATATGAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACAACTTT 2419
Qy 2062 CTACCATCTTGTAGCTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGTTGCTGTCAT 2121
Db 2420 CTACCATCTTGTAGCTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGTTGCTGTCAT 2479
Qy 2122 CGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAACGAGTTAGCAATCAGATCTT 2181
Db 2480 CGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAACGAGTTAGCAATCAGATCTT 2539

Qy 2182 CTTTCAGGTTATGACCATCTGTGCAGTTTCGTAATGTCGTCAACTTTCCGACTCTGAG 2241
Db 2540 CTTTCAGGTTATGACCATCTGTGCAGTTTCGTAATGTCGTCAACTTTCCGACTCTGAG 2599
Qy 2242 AAATCTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGACAGAACAC 2301
Db 2600 AAATCTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGACAGAACAC 2659
Qy 2302 GGATATATAGTGGATGTGTCAAAACGCATACCATTTTGAAT 2342
Db 2660 GGATATATAGTGGATGTGTCAAAACGCATACCATTTTGAAT 2700
RESULT 3
US-10-369-493-46942
; Sequence 46942, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46942
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-10-369-493-46942

Query Match 34.3%; Score 1081; DB 6; Length 1170;
Best Local Similarity 100.0%; Pred. No. 3.9e-461;
Matches 1081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22 TTGAAGATTGATTCCTCGTTAAACGAGCGGTTAGACAGAAATGAAAGAACGCGGTACAT 81
Db 1 TTGAAGATTGATTCCTCGTTAAACGAGCGGTTAGACAGAAATGAAAGAACGCGGTACAT 60
Qy 82 CGTAACCTGCGGTCAATGGATGAGCGCGGTTCCAGAGAGAAATATTTGATGGCGAAAT 141
Db 61 CGTAACCTGCGGTCAATGGATGAGCGCGGTTCCAGAGAGAAATATTTGATGGCGAAAT 120
Qy 142 CAACGGTCTGCTCTCAAAACAATTTTAGGGCTCGCAAGCGATAGACGTTTCATCGAT 201
Db 121 CAACGGTCTGCTCTCAAAACAATTTTAGGGCTCGCAAGCGATAGACGTTTCATCGAT 180
Qy 202 CGAGCCCAACAGCATTCGACCAATTTGGGACAGGAGCAGCGGTTACGTTTAAACGACA 261
Db 181 CGAGCCCAACAGCATTCGACCAATTTGGGACAGGAGCAGCGGTTACGTTTAAACGACA 240
Qy 262 GGCAATTCGGTCTGGCATGAAAGCTAGAAAGAGATTTGCCGTTTAAACTGACAGAA 321
Db 241 GGCAATTCGGTCTGGCATGAAAGCTAGAAAGAGATTTGCCGTTTAAACTGACAGAA 300
Qy 322 CGGCGCTGCTGTTTTCGAGCGGTTACTTTGGCCCAATGTCGGTGTCTTCATCTGCCA 381
Db 301 CGGCGCTGCTGTTTTCGAGCGGTTACTTTGGCCCAATGTCGGTGTCTTCATCTGCCA 360
Qy 382 GAAAGGAGAGATGTCATTTTAAGTGACCGATCTCAATCATGCAAGTATGATCGAGGCTGC 441
Db 361 GAAAGGAGAGATGTCATTTTAAGTGACCGATCTCAATCATGCAAGTATGATCGAGGCTGC 420
Qy 442 CGACTTTCTAAGGCTGATACAGTTGTTTATCGGCAATTTGATGATGATGATGATGATGAT 501
Db 421 CGACTTTCTAAGGCTGATACAGTTGTTTATCGGCAATTTGATGATGATGATGATGATGAT 480

Qy 502 AAGCTGAATGAACACACAGCGTTATCAGCGCGTTTTATCGTAACACAGCGAGTATTGAGC 561
Db 481 AAGCTGAATGAACACACAGCGTTATCAGCGCGTTTTATCGTAACACAGCGAGTATTGAGC 540
Qy 562 ATGGATGGCACAATCGCCCTCTTTGATCAGATCATCTCACTTGTGAAACGCTATCATGCC 621
Db 541 ATGGATGGCACAATCGCCCTCTTTGATCAGATCATCTCACTTGTGAAACGCTATCATGCC 600
Qy 622 TTCTGTGTCGTTGATGATGCCAGCAACAGGAGTTTGGCGGATTTGGGACAGGAAACG 681
Db 601 TTCTGTGTCGTTGATGATGCCAGCAACAGGAGTTTGGCGGATTTGGGACAGGAAACG 660
Qy 682 AGTGAATACCTTTGGTGTTCGCGACATTTGATCGGCACCTTAAGCAAAAGCTGTGGC 741
Db 661 AGTGAATACCTTTGGTGTTCGCGACATTTGATCGGCACCTTAAGCAAAAGCTGTGGC 720
Qy 742 GCGGAAGGAGGTTTTCGCGAGGATCAGCGGTCTTCATCGACTTTTGTGTAACCATGCC 801
Db 721 GCGGAAGGAGGTTTTCGCGAGGATCAGCGGTCTTCATCGACTTTTGTGTAACCATGCC 780
Qy 802 AGAATATTTATCTTTGAAACCGCTATTCCGCCAGCAGCTGTGCGGTGCTCAGAGGCT 861
Db 781 AGAATATTTATCTTTGAAACCGCTATTCCGCCAGCAGCTGTGCGGTGCTCAGAGGCT 840
Qy 862 TTCAACATCATTTGAAGCCAGCAGGAAACACAGAGCTTTATTTCTTATATCAGCATG 921
Db 841 TTCAACATCATTTGAAGCCAGCAGGAAACACAGAGCTTTATTTCTTATATCAGCATG 900
Qy 922 ATCAGAACCATGCTGAAGAAATATGGGTTATGTGTGAAAGAGATCACAACCGATTAT 981
Db 901 ATCAGAACCATGCTGAAGAAATATGGGTTATGTGTGAAAGAGATCACAACCGATTAT 960
Qy 982 CCTGTAGTCATTGCGGATGCCATATAAACCGTCTCTATTTCGTGAAAACTGCAGGGCAAG 1041
Db 961 CCTGTAGTCATTGCGGATGCCATATAAACCGTCTCTATTTCGTGAAAACTGCAGGGCAAG 1020
Qy 1042 GGAATTTATGCTCTGCGCATTTCCGCCCGCAACCGTTCGCGCGGTGAAAGCCGATTGCA 1101
Db 1021 GGAATTTATGCTCTGCGCATTTCCGCCCGCAACCGTTCGCGCGGTGAAAGCCGATTGCA 1080
Qy 1102 A 1102
Db 1081 A 1081

RESULT 4
US-09-869-855A-3
; Sequence 3, Application US/09869855A
; Publication No. US20030044940A1
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Method For Increasing Gene Copy Number
; FILE REFERENCE: 10028.204-US
; CURRENT APPLICATION NUMBER: US/09/869,855A
; CURRENT FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5793
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-869-855A-3

Query Match 27.5%; Score 867; DB 3; Length 5793;
Best Local Similarity 100.0%; Pred. No. 7.4e-207;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1475 TGCTTAGGAAGACAGTTTATTAAGTGAATGAAGACGGTGTCTCTCAATATTTCTTAT 1534
Db 968 TGCTTAGGAAGACAGTTTATTAAGTGAATGAAGACGGTGTCTCTCAATATTTCTTAT 1027

Qy 1535 TTAGAAAGCAATCTAAAATTATCTGAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1594
Db 1028 TTAGAAAGCAATCTAAAATTATCTGAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1087
Qy 1595 TAATGACTAGAGAAAGAAATGAAGATCTTCAATGAATTTAGGAACGAATATTCGATA 1654
Db 1088 TAATGACTAGAGAAAGAAATGAAGATCTTCAATGAATTTAGGAACGAATATTCGATA 1147
Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTTGGTCTGTCAGACTGATG 1714
Db 1148 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTTGGTCTGTCAGACTGATG 1207
Qy 1715 GGCCCTATTTCGATATTTGAGATGATGTCATGTCATCAACAGAGGAAGAGTTCAGCC 1774
Db 1208 GGCCCTATTTCGATATTTGAGATGATGTCATGTCATCAACAGAGGAAGAGTTCAGCC 1267
Qy 1775 ATGAATGGACCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGCGAAGAGATTCCTAC 1834
Db 1268 ATGAATGGACCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGCGAAGAGATTCCTAC 1327
Qy 1835 TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
Db 1328 TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1387
Qy 1895 TTTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db 1388 TTTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1447
Qy 1955 TAGAGGCCCAACGTTCCAGATGCGAATTTTGCCCTTATCGTAGAGAGCTGTTGAAT 2014
Db 1448 TAGAGGCCCAACGTTCCAGATGCGAATTTTGCCCTTATCGTAGAGAGCTGTTGAAT 1507
Qy 2015 ATGCAGCAAAATGCGGTAAATATTCGTGTGCAAGGACCGACAAATTTCTACCATCTCTGA 2074
Db 1508 ATGCAGCAAAATGCGGTAAATATTCGTGTGCAAGGACCGACAAATTTCTACCATCTCTGA 1567
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTCATCGCATCGCATCTGTTATA 2134
Db 1568 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTCATCGCATCGCATCTGTTATA 1627
Qy 2135 CGAGGAGCGCTTCGGTCTTAACCTGAACAGTAAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db 1628 CGAGGAGCGCTTCGGTCTTAACCTGAACAGTAAAGCAATCAGATCTTCCTTCAGGTTATG 1687
Qy 2195 ACCATCTGTCCAGTTCGTAAATGTCGTCAACTTTCCGACTCTGAGAACTTCCTGGAAT 2254
Db 1688 ACCATCTGTCCAGTTCGTAAATGTCGTCAACTTTCCGACTCTGAGAACTTCCTGGAAT 1747
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACCGGATATATAGTGG 2314
Db 1748 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACCGGATATATAGTGG 1807
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGA 2341
Db 1808 ATGTGTCAAAACGCATACCAATTTTGA 1834

RESULT 5
US-10-402-466A-2
; Sequence 2, Application US/10402466A
; Publication No. US20040028695A1
; GENERAL INFORMATION:
; APPLICANT: Park, Sukjoon
; APPLICANT: Giri, Lallan
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING
; TITLE OF INVENTION: AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS
; FILE REFERENCE: 18933-00005
; CURRENT APPLICATION NUMBER: US/10/402,466A
; PRIOR FILING DATE: 2003-03-20
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 2

; LENGTH: 5865
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION: Artificial DNA sequence to be used as one of the two backbones of
; OTHER INFORMATION: E. coli-Bacillus anthracis shuttle vectors. Designated as pBP II
US-10-402-466A-2
Query Match 27.5%; Score 867; DB 7; Length 5865;
Best Local Similarity 100.0%; Pred. No. 7.4e-207; Indels 0; Gaps 0;
Matches 867; Conservative 0; Mismatches 0;
Qy 1475 TGCTTAGGAAGACGAGTTAATAATAGCTGAATTAAGAACGGTCTCTCCAAATATTCCTAT 1534
Db 4390 TGCTTAGGAAGACGAGTTAATAATAGCTGAATTAAGAACGGTCTCTCCAAATATTCCTAT 4449
Qy 1535 TTAGAAAAGCAATCTAAAATTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594
Db 4450 TTAGAAAAGCAATCTAAAATTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 4509
Qy 1595 TAATGACTAGAGAAAGAAATGAAGATTTCTCATGAATTTAAGGAACGAATATTCGATA 1654
Db 4510 TAATGACTAGAGAAAGAAATGAAGATTTCTCATGAATTTAAGGAACGAATATTCGATA 4569
Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTTGGTCTGTCAGACTGATG 1714
Db 4570 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTTGGTCTGTCAGACTGATG 4629
Qy 1715 GGCCCTATTTCGATATTTGAGATGATGTCATGTCATCAACAGAGGAAGAGTTCAGCC 1774
Db 4630 GGCCCTATTTCGATATTTGAGATGATGTCATGTCATCAACAGAGGAAGAGTTCAGCC 4689
Qy 1775 ATCAATGGACCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGCGAAGAGATTCCTAC 1834
Db 4690 ATCAATGGACCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGCGAAGAGATTCCTAC 4749
Qy 1835 TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
Db 4750 TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 4809
Qy 1895 TTTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db 4810 TTTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 4869
Qy 1955 TAGAGGCCCAACGTTCCACGATGCGAATTTGTGCCCTTATCGTAGAGAGCTGTTGAAT 2014
Db 4870 TAGAGGCCCAACGTTCCACGATGCGAATTTGTGCCCTTATCGTAGAGAGCTGTTGAAT 4929
Qy 2015 ATGCAGCAAAATGGCGTAAATATTCGTGTGCAAGGACCGACAAATTTCTACCATCTCTGA 2074
Db 4930 ATGCAGCAAAATGGCGTAAATATTCGTGTGCAAGGACCGACAAATTTCTACCATCTCTGA 4989
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTCATCGCATCGCATCTGTTATA 2134
Db 4990 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTCATCGCATCGCATCTGTTATA 5049
Qy 2135 CGAGGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db 5050 CGAGGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 5109
Qy 2195 ACCATCTGTCCAGTTCGTAAATGTCGTGGTCAACTTTCCGACTCTGAGAACTTCCTGGAAT 2254
Db 5110 ACCATCTGTCCAGTTCGTAAATGTCGTGGTCAACTTTCCGACTCTGAGAACTTCCTGGAAT 5169
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACCGGATATATAGTGG 2314
Db 5170 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACCGGATATATAGTGG 5229
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGA 2341
Db 5230 ATGTGTCAAAACGCATACCAATTTTGA 5256

RESULT 6

US-09-859-855A-2

; Sequence 2, Application US/09869855A

; Publication No. US20030044940A1

; GENERAL INFORMATION:

; APPLICANT: Rasmussen, Michael Dolberg

; TITLE OF INVENTION: Method For Increasing Gene Copy Number

; FILE REFERENCE: 10028.204-US

; CURRENT APPLICATION NUMBER: US/09/869,855A

; CURRENT FILING DATE: 2001-07-05

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 5943

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

US-09-859-855A-2

Query Match 27.5%; Score 867; DB 3; Length 5943;

Best Local Similarity 100.0%; Pred. No. 7.5e-207;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1475	TGCTTAGGAACGAGTTATTAAAGCTGGAATAAGAACGGTGCTCTCCAAATATCTTAT	1534
Db	968	TGCTTAGGAACGAGTTATTAAAGCTGGAATAAGAACGGTGCTCTCCAAATATCTTAT	1027
Qy	1535	TTAGAAAAGCAAAATCTTAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	1594
Db	1028	TTAGAAAAGCAAAATCTTAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	1087
Qy	1595	TAATGACTAGAGAAGAAAGATGAAGATGTTCATGAAATTAAGAAAGCAATATTTGGATA	1654
Db	1088	TAATGACTAGAGAAGAAAGATGAAGATGTTCATGAAATTAAGAAAGCAATATTTGGATA	1147
Qy	1655	AATATGGGATGATGTTAAGGCTATGGTGTTTATGGCTCTCTGGTCGTCAGACTGATG	1714
Db	1148	AATATGGGATGATGTTAAGGCTATGGTGTTTATGGCTCTCTGGTCGTCAGACTGATG	1207
Qy	1715	GGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCAACAGAGAAAGCAGATTTCAGCC	1774
Db	1208	GGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCAACAGAGAAAGCAGATTTCAGCC	1267
Qy	1775	ATGAATGACAAACCGGTGAGTGGAAAGTGAATTTTGTATAGCAAGAGATTCTTAC	1834
Db	1268	ATGAATGACAAACCGGTGAGTGGAAAGTGAATTTTGTATAGCAAGAGATTCTTAC	1327
Qy	1835	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1894
Db	1328	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1387
Qy	1895	TTTTGCCGATTTATGATTCAGGTGGATACTTAGAGAAAGTGTATCAAACTGCTAAATCGG	1954
Db	1388	TTTTGCCGATTTATGATTCAGGTGGATACTTAGAGAAAGTGTATCAAACTGCTAAATCGG	1447
Qy	1955	TAGAAGCCAAACGGTTCCAGATGGATTTGTGGCTTATCGTAGAGAGCTGTTGAAAT	2014
Db	1448	TAGAAGCCAAACGGTTCCAGATGGATTTGTGGCTTATCGTAGAGAGCTGTTGAAAT	1507
Qy	2015	ATGACAGGCAAAATGGGTAATATTCGTGTGCAAGGACCGACAACATTTCTACCATCTTGA	2074
Db	1508	ATGACAGGCAAAATGGGTAATATTCGTGTGCAAGGACCGACAACATTTCTACCATCTTGA	1567
Qy	2075	CTGTACAGGTAGCAATGGCAGGTGCATGTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTG	2134
Db	1568	CTGTACAGGTAGCAATGGCAGGTGCATGTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTG	1627
Qy	2135	CGACAGGCGCTTCGGTCTTAACTAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG	2194
Db	1628	CGACAGGCGCTTCGGTCTTAACTAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG	1687

Qy	2195	ACCATCTGTGCCAGTTCGTAAATGTCTGCTCAACTTTCCGACTCTGAGAAACTTCTGGAAT	2254
Db	1688	ACCATCTGTGCCAGTTCGTAAATGTCTGCTCAACTTTCCGACTCTGAGAAACTTCTGGAAT	1747
Qy	2255	CGCTAGAGAAATTTCTGGAATGGGAATTCAGAGTGGACAGAACGACAGGATATATAGTGG	2314
Db	1748	CGCTAGAGAAATTTCTGGAATGGGAATTCAGAGTGGACAGAACGACAGGATATATAGTGG	1807
Qy	2315	ATGTGTCAAAACGCATACCATTTTGA 2341	
Db	1808	ATGTGTCAAAACGCATACCATTTTGA 1834	

RESULT 7

US-10-402-466A-1

; Sequence 1, Application US/10402466A

; Publication No. US20040028695A1

; GENERAL INFORMATION:

; APPLICANT: Park, Sukjoon

; APPLICANT: Giri, Lallan

; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING

; AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS

; FILE REFERENCE: 18933-00005

; CURRENT APPLICATION NUMBER: US/10/402,466A

; CURRENT FILING DATE: 2003-03-20

; PRIOR APPLICATION NUMBER: 60/372,152

; PRIOR FILING DATE: 2002-04-12

; NUMBER OF SEQ ID NOS: 34

; SEQ ID NO 1

; LENGTH: 6694

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; NAME/KEY:

; LOCATION:

; OTHER INFORMATION: Artificial DNA sequence to be used as one of the two backbones of

US-10-402-466A-1

E. coli-Bacillus anthracis shuttle vectors. Designated as pBP 1.

Query Match 27.5%; Score 867; DB 7; Length 6694;

Best Local Similarity 100.0%; Pred. No. 8e-207;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1475	TGCTTAGGAACGAGTTATTAAATAGCTGAATAAGAACGGTGCTCTCCAAATATCTTAT	1534
Db	5219	TGCTTAGGAACGAGTTATTAAATAGCTGAATAAGAACGGTGCTCTCCAAATATCTTAT	5278
Qy	1535	TTAGAAAAGCAAAATCTTAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	1594
Db	5279	TTAGAAAAGCAAAATCTTAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	5338
Qy	1595	TAATGACTAGAGAAGAAAGATGAGATTGTTCTATGAAATTAAGGAACGAATATTTGGATA	1654
Db	5339	TAATGACTAGAGAAGAAAGATGAGATTGTTCTATGAAATTAAGGAACGAATATTTGGATA	5398
Qy	1655	AAATATGGGATGATGTTTAAAGCTATTTCGTGTTTATGCTCTCTTTGGTCGTGCTGACTGATG	1714
Db	5399	AAATATGGGATGATGTTTAAAGCTATTTCGTGTTTATGCTCTCTTTGGTCGTGCTGACTGATG	5458
Qy	1715	GGCCCTATTCCGATATTGAGATGATGTGTGTCATGTCAACAGAGAAAGCAGATTTCAGCC	1774
Db	5459	GGCCCTATTCCGATATTGAGATGATGTGTGTCATGTCAACAGAGAAAGCAGATTTCAGCC	5518
Qy	1775	ATGAATGGACAAACCGGTGAGTGGAAAGTGAATTTTGTAGCGAAGAGATTCTTAC	1834
Db	5519	ATGAATGGACAAACCGGTGAGTGGAAAGTGAATTTTGTAGCGAAGAGATTCTTAC	5578
Qy	1835	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1894
Db	5579	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	5638
Qy	1895	TTTTGCCGATTTATGATTCAGGTGGATACTTAGAGAAAGTGTATCAAACTGCTAAATCGG	1954

Db 5639 TTTTGGCGATTTATGATTCAGGTGGATACCTAGAGAAAGTGATCAAACCTGCTAAATCGG 5698
Qy 1955 TAGAAGCCCAACGCTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014
Db 5699 TAGAAGCCCAACGCTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 5758
Qy 2015 ATGACGCAAAATGCGCTTAATTTGCTGTGCAAGGACCGACAACATTTCTACCATCTTGA 2074
Db 5759 ATGACGCAAAATGCGCTTAATTTGCTGTGCAAGGACCGACAACATTTCTACCATCTTGA 5818
Qy 2075 CTGTACAGGTAGCAATGCGAGTGCATGTTGATGCTGTGCATCATCGCATCTGTTATA 2134
Db 5819 CTGTACAGGTAGCAATGCGAGTGCATGTTGATGCTGTGCATCATCGCATCTGTTATA 5878
Qy 2135 CGACGAGCGCTTCGGTCTTAACGAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db 5879 CGACGAGCGCTTCGGTCTTAACGAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 5938
Qy 2195 ACCATCTGTCCAGTTCGTAATGCTGTGCAACCTTTCCGACTCTGAGAACTTCTGGAAT 2254
Db 5939 ACCATCTGTCCAGTTCGTAATGCTGTGCAACCTTTCCGACTCTGAGAACTTCTGGAAT 5998
Qy 2255 CGCTAGAGATTTCTGAAATGGGATTCAGAGTGGACAGAACGACGAGATATATAGTGG 2314
Db 5999 CGCTAGAGATTTCTGAAATGGGATTCAGAGTGGACAGAACGACGAGATATATAGTGG 6058
Qy 2315 ATGTGTCAAAACGCATACCATTTTGA 2341
Db 6059 ATGTGTCAAAACGCATACCATTTTGA 6085

RESULT 8

US-10-084-814-3
; Sequence 3, Application US/10084814
; Publication No. US20030108982A1
; GENERAL INFORMATION:
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN,
; ERIC BASTIAAN
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF
; STEROIDS AND GENETICALLY ENGINEERED CELLS
; USED THEREIN
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN, MUSELIAN & LUCAS
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,814
; FILING DATE: 26-Feb-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/418,085
; FILING DATE: 06-APR-1995
; APPLICATION NUMBER: 08/054,185
; FILING DATE: 26-APR-1993
; APPLICATION NUMBER: 08/002,608
; FILING DATE: 11-JAN-1993
; APPLICATION NUMBER: 07/474,857
; FILING DATE: 30-OCT-1990
; APPLICATION NUMBER: 07/474,798
; FILING DATE: 16-JULY-1990
; APPLICATION NUMBER: PCT/NL89/00072
; FILING DATE: 25-SEPT-1989
; APPLICATION NUMBER: NL88/200904.6

; FILING DATE: 06-MAY-1988
; APPLICATION NUMBER: NL/88/202080.3
; FILING DATE: 03-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSELIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146.1169-
; CON-1-DIV-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7336 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
; FEATURE:
; OTHER INFORMATION: PLASMID pBHA-1
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-084-814-3
Query Match 27.5%; Score 867; DB 5; Length 7336;
Best Local Similarity 100.0%; Pred. No. 8.4e-207; Indels 0; Gaps 0;
Matches 867; Conservative 0; Mismatches 0;
Qy 1475 TGCTTAGGAGACGAGTTATTAATAGCTGAATAGAACGCTGCTCTCCAAATATTCCTAT 1534
Db 5231 TGCTTAGGAGACGAGTTATTAATAGCTGAATAGAACGCTGCTCTCCAAATATTCCTAT 5290
Qy 1535 TTAGAAAAGCAAAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594
Db 5291 TTAGAAAAGCAAAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 5350
Qy 1595 TAATGACTAGAGAGAGAAAGATCAAGATTTGTCATGAAATTAAGGACGAATATTTGATA 1654
Db 5351 TAATGACTAGAGAGAGAAAGATCAAGATTTGTCATGAAATTAAGGACGAATATTTGATA 5410
Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTCTTTATGGCTCTCTTGGTCTGAGACTGATG 1714
Db 5411 AATATGGGATGATGTTAAGGCTATTGGTCTTTATGGCTCTCTTGGTCTGAGACTGATG 5470
Qy 1715 GGCCCTATTCCGATATTGAGATGATGTCATGTCAATGTCAGAGAGAGAGTTTCAGCC 1774
Db 5471 GGCCCTATTCCGATATTGAGATGATGTCATGTCAATGTCAGAGAGAGAGTTTCAGCC 5530
Qy 1775 ATGAATGGACAACCGGTGAGTGGAAAGTGAATTTTGTAGCGAGAGAGATCTTAC 1834
Db 5531 ATGAATGGACAACCGGTGAGTGGAAAGTGAATTTTGTAGCGAGAGAGATCTTAC 5590
Qy 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCGGCTTACACATGTTGCTCAATTTTCTCTA 1894
Db 5591 TAGATTATGCATCTCAGGTGGAATCAGATTGGCGGCTTACACATGTTGCTCAATTTTCTCTA 5650
Qy 1895 TTTTGGCGATTTATGATTCAGGTGGATATTTAGAGAAAGTGTATCAAACCTGCTAAATCGG 1954
Db 5651 TTTTGGCGATTTATGATTCAGGTGGATATTTAGAGAAAGTGTATCAAACCTGCTAAATCGG 5710
Qy 1955 TAGAAGCCCAACGCTTCACGATGCGATTTGTGCCCTTATCGTAGAGAGCTGTTGAT 2014
Db 5711 TAGAAGCCCAACGCTTCACGATGCGATTTGTGCCCTTATCGTAGAGAGCTGTTGAT 5770
Qy 2015 ATGACGCAAAATGCGCTTAATTTGCTGTGCAAGGACCGACAACATTTCTACCATCTTGA 2074
Db 5771 ATGACGCAAAATGCGCTTAATTTGCTGTGCAAGGACCGACAACATTTCTACCATCTTGA 5830
Qy 2075 CTGTACAGGTAGCAATGCGAGTGCATGTTGATGCTGTGCATCATCGCATCTGTTATA 2134
Db 5831 CTGTACAGGTAGCAATGCGAGTGCATGTTGATGCTGTGCATCATCGCATCTGTTATA 5890
Qy 2135 CGACGAGCGCTTCGGTCTTAACGAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db 5891 CGACGAGCGCTTCGGTCTTAACGAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 5950

Qy 2195 ACCATCTGTCCAGTTGCTAATGCTGCTCAACTTTCCGACTCTCAGAAACTTCTGGAAT 2254
Db |||||||
Qy 5951 ACCATCTGTCCAGTTGCTAATGCTGCTCAACTTTCCGACTCTCAGAAACTTCTGGAAT 6010
Db |||||||
Qy 2255 CGTAGAGAAATTTCTGGAATGGATTTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314
Db |||||||
Qy 6011 CGTAGAGAAATTTCTGGAATGGATTTCAGGAGTGGACAGAACGACACGGATATATAGTGG 6070
Db |||||||
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGA 2341
Db 6071 ATGTGTCAAAACGCATACCAATTTTGA 6097
|||
RESULT 9
US-10-442-013-4
; Sequence 4, Application US/10442013
; Publication No. US20040014175A1
; GENERAL INFORMATION:
; APPLICANT: van der Laan, Johannes Cornelis
; APPLICANT: van Eekelen, Christiaan Albertus Gerardus
; TITLE OF INVENTION: Efficient Production of Mutant Proteases
; FILE REFERENCE: GCX329-US-C1
; CURRENT APPLICATION NUMBER: US/10/442,013
; PRIORITY FILING DATE: 2003-05-19
; PRIOR FILING DATE: 1989-08-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 7336
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pBHA-1
US-10-442-013-4

Query Match 27.5%; Score 867; DB 7; Length 7336;
Best Local Similarity 100.0%; Pred. No. 8.4e-207;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1475 TGCTTAGGAACGAGTTATTAATAGCTGAAATAAGAACGGTCTCTCAAAATATCTTAT 1534
Db |||||||
Qy 5231 TGCTTAGGAACGAGTTATTAATAGCTGAAATAAGAACGGTCTCTCAAAATATCTTAT 5290
Db |||||||
Qy 1535 TTAGAAAAGCAAACTTAAATATCTGAAAAGGAATGAGATAGTAATGGACCAATAA 1594
Db |||||||
Qy 5291 TTAGAAAAGCAAACTTAAATATCTGAAAAGGAATGAGATAGTAATGGACCAATAA 5350
Db |||||||
Qy 1595 TAATGACTAGAGAAAGAAATGAAGATTGTTCAATGAATTAAGAAAGAAATATGGATA 1654
Db |||||||
Qy 5351 TAATGACTAGAGAAAGAAATGAAGATTGTTCAATGAATTAAGAAAGAAATATGGATA 5410
Db |||||||
Qy 1655 AATATGGGATGATGTTAAGGCTATGCTGTTTATGCTCTCTTGGTCTCAGACTGATG 1714
Db |||||||
Qy 5411 AATATGGGATGATGTTAAGGCTATGCTGTTTATGCTCTCTTGGTCTCAGACTGATG 5470
Db |||||||
Qy 1715 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAGAGACGAGTTACGCC 1774
Db |||||||
Qy 5471 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAGAGAGACGAGTTACGCC 5530
Db |||||||
Qy 1775 ATGAATGGACAAACCGGTGAGTGAAGGTTGAAGTGAATTTTGTAGCGAAGATTTCTAC 1834
Db |||||||
Qy 5531 ATGAATGGACAAACCGGTGAGTGAAGGTTGAAGTGAATTTTGTAGCGAAGATTTCTAC 5590
Db |||||||
Qy 1835 TAGATTATGCACTCAGGTGGAATCAGATTGGCCGCTTACATGGTCAATTTTCTCTA 1894
Db |||||||
Qy 5591 TAGATTATGCACTCAGGTGGAATCAGATTGGCCGCTTACATGGTCAATTTTCTCTA 5650
Db |||||||
Qy 1895 TTTTGGCGATTTATGATTTCAGGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db |||||||
Qy 5651 TTTTGGCGATTTATGATTTCAGGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 5710
Db |||||||

Qy 1955 TAGAAGCCCAAAAGCTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014
Db |||||||
Qy 5711 TAGAAGCCCAAAAGCTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 5770
Db |||||||
Qy 2015 ATGAGGCAAAATGCGGTAATATTCTGCTGCAAGGACCGAACCAATTTCTACCATCTTGA 2074
Db |||||||
Qy 5771 ATGAGGCAAAATGCGGTAATATTCTGCTGCAAGGACCGAACCAATTTCTACCATCTTGA 5830
Db |||||||
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGCTGTCATCATCGCATCTGTTATA 2134
Db |||||||
Qy 5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGCTGTCATCATCGCATCTGTTATA 5890
Db |||||||
Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGTATTG 2194
Db |||||||
Qy 5891 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGTATTG 5950
Db |||||||
Qy 2195 ACCATCTGTCCAGTTGCTAATGCTGCTCAACTTTCCGACTCTTGAGAAACTTCTGGAAT 2254
Db |||||||
Qy 5951 ACCATCTGTCCAGTTGCTAATGCTGCTCAACTTTCCGACTCTTGAGAAACTTCTGGAAT 6010
Db |||||||
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314
Db |||||||
Qy 6011 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 6070
Db |||||||
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGA 2341
Db 6071 ATGTGTCAAAACGCATACCAATTTTGA 6097
|||
RESULT 10
US-10-462-128-3
; Sequence 3, Application US/10462128
; Publication No. US20040067579A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND
; GENETICALLY ENGINEERED CELLS USED THEREIN
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/462,128
; FILING DATE: 16-Jun-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/054,185
; FILING DATE: 26-APR-1993
; APPLICATION NUMBER: US/08/002,608
; FILING DATE: 11-JAN-1993
; APPLICATION NUMBER: US/07/474,857
; FILING DATE: 30-OCT-1990
; APPLICATION NUMBER: US/07/474,798
; FILING DATE: 16-JULY-1990
; APPLICATION NUMBER: PCT/NL89/00072
; FILING DATE: 25-SEPT-1989
; APPLICATION NUMBER: NL/88/200904.6
; FILING DATE: 06-MAY-88
; APPLICATION NUMBER: NL/88/202080.3
; FILING DATE: 03-SEP-88
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146.1169 CON-1
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 661-8000
TELEFAX: (212) 661-8002
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 7336
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: UNKNOWN
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: PLASMID pBHA-1
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-462-128-3

Query Match 27.5%; Score 867; DB 7; Length 7336;
Best Local Similarity 100.0%; Pred. No. 8.4e-207;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1475 TGCCTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTTAT 1534
DB 5231 TGCCTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTTAT 5290
QY 1535 TTAGAAAAGCAATCTAAATATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594
DB 5291 TTAGAAAAGCAATCTAAATATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 5350
QY 1595 TAATGACTAGAGAAGAAAGTAAGATTTGTTTCATGAAATTAAGGAACGAATATTGGATA 1654
DB 5351 TAATGACTAGAGAAGAAAGTAAGATTTGTTTCATGAAATTAAGGAACGAATATTGGATA 5410
QY 1655 AATATGGGATGATGTTAAAGCTATTGTTGTTTATGGCTCTCTTTGGTGGTGCAGACTGATG 1714
DB 5411 AATATGGGATGATGTTAAAGCTATTGTTGTTTATGGCTCTCTTTGGTGGTGCAGACTGATG 5470
QY 1715 GGCCCTATTTCGGATATTGAGATGATGTGTCATGTCAATGCAACAGAGGAAGCAGAGTTTCAGCC 1774
DB 5471 GGCCCTATTTCGGATATTGAGATGATGTGTCATGTCAACAGAGGAAGCAGAGTTTCAGCC 5530
QY 1775 ATGAATGGACAACCGGTGAGTGGAAAGTGAAGTGAATTTTGTATAGCGAAGAGATTCTAC 1834
DB 5531 ATGAATGGACAACCGGTGAGTGGAAAGTGAAGTGAATTTTGTATAGCGAAGAGATTCTAC 5590
QY 1835 TAGATTATGATCTCAGTGGAAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
DB 5591 TAGATTATGATCTCAGTGGAAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 5650
QY 1895 TTTTGGCGATTATGATTTCAGGTGGATACCTTAGAGAAGTGTATCAAACTGCTAAATCGG 1954
DB 5651 TTTTGGCGATTATGATTTCAGGTGGATACCTTAGAGAAGTGTATCAAACTGCTAAATCGG 5710
QY 1955 TAGAAGCCAAAGCTTCACGATGCGAATTTGTGCGCTTATTCGTAGAAAGAGCTGTTGAAT 2014
DB 5711 TAGAAGCCAAAGCTTCACGATGCGAATTTGTGCGCTTATTCGTAGAAAGAGCTGTTGAAT 5770
QY 2015 ATCAGGCAATCGCGTAATATTCGTGTGCAAGCGACGCAACATTTCTACCACTCCTTGA 2074
DB 5771 ATCAGGCAATCGCGTAATATTCGTGTGCAAGCGACGCAACATTTCTACCACTCCTTGA 5830
QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGTATGTTGCTGTGCATCATCGCATCTGTTATA 2134
DB 5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGTATGTTGCTGTGCATCATCGCATCTGTTATA 5890
QY 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
DB 5891 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 5950
QY 2195 ACCATCTGTCCAGTTCGTAATGTCGTCAACTTTTCCGACTCTGAGAACTCTTCGGAAT 2254
DB 5951 ACCATCTGTCCAGTTCGTAATGTCGTCAACTTTTCCGACTCTGAGAACTCTTCGGAAT 6010
QY 2255 CGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGGACAGAACGACCGGATATATAGTGG 2314

DB 6011 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACACGCGATATATAGTGG 6070
QY 2315 ATGTGTCAAAACGCGATACCATTTTGA 2341
DB 6071 ATGTGTCAAAACGCGATACCATTTTGA 6097
RESULT 11
US-10-402-466A-7
; Sequence 7, Application US/10402466A
; Publication No. US20040028695A1
; GENERAL INFORMATION:
; APPLICANT: Park, Sukjoon
; APPLICANT: Giri, Lallan
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING
; FILE REFERENCE: 18933-00005
; CURRENT APPLICATION NUMBER: US/10/402,466A
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: 60/372,152
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 7
; LENGTH: 8198
; TYPE: DNA
; ORGANISM: Bacillus anthracis
; FEATURES:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION: Expression vector for B. anthracis Protective Antigen (PA). Enti
; OTHER INFORMATION: sequence is shown since the vector sequence is different from the
; OTHER INFORMATION: of the pBP vectors. The PA coding sequence is from 3735 to 5942.
US-10-402-466A-7
Query Match 27.5%; Score 867; DB 7; Length 8198;
Best Local Similarity 100.0%; Pred. No. 8.9e-207;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1475 TGCCTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTTAT 1534
DB 2259 TGCCTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTTAT 2318
QY 1535 TTAGAAAAGCAATCTAAATATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594
DB 2319 TTAGAAAAGCAATCTAAATATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 2378
QY 1595 TAATGACTAGAGAAGAAAGTAAGATTTGTTTCATGAAATTAAGGAACGAATATTGGATA 1654
DB 2379 TAATGACTAGAGAAGAAAGTAAGATTTGTTTCATGAAATTAAGGAACGAATATTGGATA 2438
QY 1655 AATATGGGATGATGTTAAAGCTATTGGTGTGTTTATGGCTCTCTTTGGTGGTGCAGACTGATG 1714
DB 2439 AATATGGGATGATGTTAAAGCTATTGGTGTGTTTATGGCTCTCTTTGGTGGTGCAGACTGATG 2498
QY 1715 GGCCCTATTTCGGATATTGAGATGATGTGTCATGTCAACAGAGGAAGCAGAGTTTCAGCC 1774
DB 2499 GGCCCTATTTCGGATATTGAGATGATGTGTCATGTCAACAGAGGAAGCAGAGTTTCAGCC 2558
QY 1775 ATGAATGGACAACCGGTGAGTGGAAAGTGAATTTTGTATAGCGAAGAGATTCTAC 1834
DB 2559 ATGAATGGACAACCGGTGAGTGGAAAGTGAATTTTGTATAGCGAAGAGATTCTAC 2618
QY 1835 TAGATTATGATCTCAGTGGAAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
DB 2619 TAGATTATGATCTCAGTGGAAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 2678
QY 1895 TTTTGGCGATTATGATTTCAGGTGGATACCTTAGAGAAGTGTATCAAACTGCTAAATCGG 1954
DB 2679 TTTTGGCGATTATGATTTCAGGTGGATACCTTAGAGAAGTGTATCAAACTGCTAAATCGG 2738
QY 1955 TAGAAGCCAAAGCTTCACGATGCGAATTTGTGCGCTTATTCGTAGAAAGAGCTGTTGAAT 2014

```
Db 2739 TAGAAGCCAAACGGTCCACGATGCGATTTGTGCGCCTTATCGTAGAGAGCTGTTTGAAT 2798
Qy 2015 ATGCAGGCAAAATGGGTAATATTGGTGGCAAGACCGACAACATTTCTACCATCCTTGA 2074
Db 2799 ATGCAGGCAAAATGGGTAATATTGGTGGCAAGACCGACAACATTTCTACCATCCTTGA 2858
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA 2134
Db 2859 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA 2918
Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db 2919 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2978
Qy 2195 ACCATCTGTGCCAGTTCGTAATGCTGTCACTTCCGACTCTGAGAACTTCTGGAAAT 2254
Db 2979 ACCATCTGTGCCAGTTCGTAATGCTGTCACTTCCGACTCTGAGAACTTCTGGAAAT 3038
Qy 2255 CGCTAGAGAAATTTCTGGAAATGGGATTACAGGAGTGGACAGACGACACGGATATATAGTGG 2314
Db 3039 CGCTAGAGAAATTTCTGGAAATGGGATTACAGGAGTGGACAGACGACACGGATATATAGTGG 3098
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGAA 2341
Db 3099 ATGTGTCAAAACGCATACCAATTTTGAA 3125

RESULT 12
US-10-402-466A-10
; Sequence 10, Application US/10402466A
; Publication No. US20040028695A1
; GENERAL INFORMATION:
; APPLICANT: Park, Sukjoon
; APPLICANT: Giri, Laljan
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING
; TITLE OF INVENTION: AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS
; FILE REFERENCE: 18933-00005
; CURRENT APPLICATION NUMBER: US/10/402,466A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: 60/372,152
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 10
; LENGTH: 9286
; TYPE: DNA
; ORGANISM: Bacillus anthracis
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION: Dual expression vector pBP105 for B. anthracis PA and LF30. Enti
; OTHER INFORMATION: sequence is shown since the vector sequence contains two coding
; OTHER INFORMATION: regions. The coding region for PA is from 3735 to 5942 and the c
; OTHER INFORMATION: region for LF30 is from 6391 to 7161.
US-10-402-466A-10

Query Match 27.5%; Score 867; DB 7; Length 9286;
Best Local Similarity 100.0%; Pred. No. 9.6e-207;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGCAGGTTTATTAATAGCTGAATAAGAACGGTGTCTCCAAATATTTCTTAT 1534
Db 2259 TGCTTAGGAAGCAGGTTTATTAATAGCTGAATAAGAACGGTGTCTCCAAATATTTCTTAT 2318
Qy 1535 TTAGAAAAGCAAAATCTAAATATCTGAAAAGGGAATGAGAAATAGTGAATGAGCAATATA 1594
Db 2319 TTAGAAAAGCAAAATCTAAATATCTGAAAAGGGAATGAGAAATAGTGAATGAGCAATATA 2378
Qy 1595 TAATGACTAGAGAGAAAGATGAAGATTGTCATGAATTAAGAAACGGAATATTTGGATA 1654
Db 2379 TAATGACTAGAGAGAAAGATGAAGATTGTCATGAATTAAGAAACGGAATATTTGGATA 2438
Qy 1655 AATATGGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTTGGTGTGCTGATG 1714
```

```
Db 2439 AATATGGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTTGGTGTGCTGAGACTGATG 2498
Qy 1715 GGCCCTATTTCGGATATTGAGATGATGTGTCTCATGTCAACAGAGGAAGCAGAGTTTCAGCC 1774
Db 2499 GGCCCTATTTCGGATATTGAGATGATGTGTCTCATGTCAACAGAGGAAGCAGAGTTTCAGCC 2558
Qy 1775 ATGAATGGACAACCGGTGAGTGGAAAGGTGGAAGTGAATTTTGTATAGCGAAGAGATTCTTAC 1834
Db 2559 ATGAATGGACAACCGGTGAGTGGAAAGGTGGAAGTGAATTTTGTATAGCGAAGAGATTCTTAC 2618
Qy 1835 TAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTACACATGTTCAATTTTCTTCTA 1894
Db 2619 TAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTACACATGTTCAATTTTCTTCTA 2678
Qy 1895 TTTTGGCCGATTTATGATTTAGATTTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db 2679 TTTTGGCCGATTTATGATTTAGATTTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 2738
Qy 1955 TAGAAGCCCAAAAGCTTCCACGATGCGATTGTTGTGCCCTTATCGTAGAAGAGAGCTGTTTGAAT 2014
Db 2739 TAGAAGCCCAAAAGCTTCCACGATGCGATTGTTGTGCCCTTATCGTAGAAGAGAGCTGTTTGAAT 2798
Qy 2015 ATGCAGGCAAAATGGCGTAATATTTCGTGTGCAAGACCGACAACATTTCTACCATCCTTGA 2074
Db 2799 ATGCAGGCAAAATGGCGTAATATTTCGTGTGCAAGACCGACAACATTTCTACCATCCTTGA 2858
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGTCATCGCATCGCATCTGTTATA 2134
Db 2859 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGTCATCGCATCGCATCTGTTATA 2918
Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194
Db 2919 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2978
Qy 2195 ACCATCTGTGCCAGTTCGTAATGCTGTGTTCAACTTCCTGCACTCTGAGAACTTCTGGAAAT 2254
Db 2979 ACCATCTGTGCCAGTTCGTAATGCTGTGTTCAACTTCCTGCACTCTGAGAACTTCTGGAAAT 3038
Qy 2255 CGCTAGAGAAATTTCTGGAAATGGGATTACAGGAGTGGACAGACGACACGGATATATAGTGG 2314
Db 3039 CGCTAGAGAAATTTCTGGAAATGGGATTACAGGAGTGGACAGACGACACGGATATATAGTGG 3098
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGAA 2341
Db 3099 ATGTGTCAAAACGCATACCAATTTTGAA 3125

RESULT 13
US-10-032-393-7/c
; Sequence 7, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITEA 010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 10929
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector pBP25
```


US-10-032-393-7

Query Match 27.5%; Score 867; DB 5; Length 10929;
Best Local Similarity 100.0%; Pred. No. 1e-206;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGCTCTCAAATATTCCTAT 1534
DB 8753 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGCTCTCAAATATTCCTAT 8694

QY 1535 TTAGAAAAGCAATCTAAATATCTGAAAAGGAATGAGATAGTGAATGGAACCAATAA 1594
DB 8693 TTAGAAAAGCAATCTAAATATCTGAAAAGGAATGAGATAGTGAATGGAACCAATAA 8634

QY 1595 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAAACGAATATGGATA 1654
DB 8633 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAAACGAATATGGATA 8574

QY 1655 AATATGGGGATGATGTTAAAGCTATTGGTGTATAGCTCTCTTGGTGGTCAGACTGATG 1714
DB 8573 AATATGGGGATGATGTTAAAGCTATTGGTGTATAGCTCTCTTGGTGGTCAGACTGATG 8514

QY 1715 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAGAGACGAGTTCAGCC 1774
DB 8513 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAGAGACGAGTTCAGCC 8454

QY 1775 ATGAATGGAACAACCGGTGAGTGAAGGTGGAAGTGAATTTTATAGCGAAGAGATTCCTAC 1834
DB 8453 ATGAATGGAACAACCGGTGAGTGAAGGTGGAAGTGAATTTTATAGCGAAGAGATTCCTAC 8394

QY 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
DB 8393 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 8334

QY 1895 TTTTGCCGATTTATGATTCAGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
DB 8333 TTTTGCCGATTTATGATTCAGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 8274

QY 1955 TAGAAGCCCAACCGTCCAGATCGGATTTGTGCCCTTATCGTAGAAGAGCTGTTCAAT 2014
DB 8273 TAGAAGCCCAACCGTCCAGATCGGATTTGTGCCCTTATCGTAGAAGAGCTGTTCAAT 8214

QY 2015 ATGACGCAAAATGCGGTAAATATTCGTGTGAAGACCGCAACAATTTCTACCACTCTGA 2074
DB 8213 ATGACGCAAAATGCGGTAAATATTCGTGTGAAGACCGCAACAATTTCTACCACTCTGA 8154

QY 2075 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2134
DB 8153 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 8094

QY 2135 CGACGAGCGCTTCGGTCTTAACTGAACGATTAAGCAATCAGATCTTCCCTTCAGGTTATG 2194
DB 8093 CGACGAGCGCTTCGGTCTTAACTGAACGATTAAGCAATCAGATCTTCCCTTCAGGTTATG 8034

QY 2195 ACCATCTGTGCGAGTTCGTATGTCGTCAACTTTTCCGACTCTGAGAAACTCTCTGGAAT 2254
DB 8033 ACCATCTGTGCGAGTTCGTATGTCGTCAACTTTTCCGACTCTGAGAAACTCTCTGGAAT 7974

QY 2255 CGCTAGAGATTTCTGGAAATGGGATTCGGAGTGGACAGAACGACAGGATATATAGTGG 2314
DB 7973 CGCTAGAGATTTCTGGAAATGGGATTCAGGAGTGGACAGAACGACAGGATATATAGTGG 7914

QY 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341
DB 7913 ATGTGTCAAAACGCATACCATTTTGAA 7887

RESULT 14

US-10-385-415-2/c
; Sequence 2, Application US/10385415
; Publication No. US20040014158A1
; GENERAL INFORMATION:
; APPLICANT: Bacher, Adelbert

APPLICANT: Fischer, Markus
TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR
PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES CONTAININ
A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES
FILE REFERENCE: 9286.6CT
CURRENT APPLICATION NUMBER: US/10/385,415
CURRENT FILING DATE: 2003-03-10
PRIOR APPLICATION NUMBER: US 09/936,028
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: PCT/EP00/01899
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: DE 19910102.7
PRIOR FILING DATE: 1999-03-08
NUMBER OF SEQ ID NOS: 154
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 5302
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: p6021-CAT Expression vector
US-10-385-415-2

Query Match 27.4%; Score 865.4; DB 7; Length 5302;
Best Local Similarity 99.9%; Pred. No. 1.8e-206;
Matches 866; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGCTCTCAAATATTCCTAT 1534
DB 3821 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGCTCTCAAATATTCCTAT 3762

QY 1535 TTAGAAAAGCAATCTAAATATCTGAAAAGGAATGAGAATAGTGAATGGAACCAATAA 1594
DB 3761 TTAGAAAAGCAATCTAAATATCTGAAAAGGAATGAGAATAGTGAATGGAACCAATAA 3702

QY 1595 TAATGACTAGAGAAGAAAGATGAGATTTGTTCAATGAAATTAAGGAAACGAATATGGATA 1654
DB 3701 TAATGACTAGAGAAGAAAGATGAGATTTGTTCAATGAAATTAAGGAAACGAATATGGATA 3642

QY 1655 AATATGGGGATGATGTTAAAGCTATTGGTGTATAGCTCTCTTGGTGGTCAGACTGATG 1714
DB 3641 AATATGGGGATGATGTTAAAGCTATTGGTGTATAGCTCTCTTGGTGGTCAGACTGATG 3582

QY 1715 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAGAGACGAGTTCAGCC 1774
DB 3581 GGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAGAGACGAGTTCAGCC 3522

QY 1775 ATGAATGGAACAACCGGTGAGTGAAGGTGGAAGTGAATTTTGTATAGCGAAGAGATTCCTAC 1834
DB 3521 ATGAATGGAACAACCGGTGAGTGAAGGTGGAAGTGAATTTTGTATAGCGAAGAGATTCCTAC 3462

QY 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894
DB 3461 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 3402

QY 1895 TTTTGCCGATTTATGATTTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
DB 3401 TTTTGCCGATTTATGATTTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 3342

QY 1955 TAGAAGCCCAACCGTCCAGATCGGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAAT 2014
DB 3341 TAGAAGCCCAACCGTCCAGATCGGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAAT 3282

QY 2015 ATGACGCAAAATGCGGTAAATATTCGTGTGAAGACCGCAACAATTTCTACCACTCTGA 2074
DB 3281 ATGACGCAAAATGCGGTAAATATTCGTGTGAAGACCGCAACAATTTCTACCACTCTGA 3222

QY 2075 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2134
DB 3221 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 3162

QY 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTGAAGCAATCAGATCTTCCCTTCAGGTTATG 2194
DB 3162 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTGAAGCAATCAGATCTTCCCTTCAGGTTATG 2194

Db 3161 CGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 3102
Qy 2195 ACCATCTGTCGACGTTCTGTAATGCTGCTCAACTTCGACTCTCGAAGAACTTCCTGGAAT 2254
Db 3101 ACCATCTGTCGACGTTCTGTAATGCTGCTCAACTTCGACTCTCGAAGAACTTCCTGGAAT 3042
Qy 2255 CGCTAGAGAAATTCCTGGAATGGGATTCAGGAGTGGACAGAACGACACCGATATATAGTGG 2314
Db 3041 CGCTAGAGAAATTCCTGGAATGGGATTCAGGAGTGGACAGAACGACACCGATATATAGTGG 2982
Qy 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341
Db 2981 ATGTGTCAAAACGCATACCATTTTGAA 2955
RESULT 15
US-10-385-415-4/c
; Sequence 4, Application US/10385415
; Publication No. US20040014158A1
; GENERAL INFORMATION:
; APPLICANT: Bacher, Adelbert
; APPLICANT: Fischer, Maribus
; TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR
; TITLE OF INVENTION: PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES CONTAININ
; TITLE OF INVENTION: A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES
; FILE REFERENCE: 9286.6CT
; CURRENT APPLICATION NUMBER: US/10/385,415
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: US 09/936,028
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/EP00/01899
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: DE 19910102.7
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 5767
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p602-BS-LuSy Expression plasmid
US-10-385-415-4
Query Match 27.4%; Score 865.4; DB 7; Length 5767;
Best Local Similarity 99.9%; Pred. No. 1.9e-206;
Matches 866; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGTCTCTCCAAATATTTCTTAT 1534
Db 4286 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGTCTCTCCAAATATTTCTTAT 4227
Qy 1535 TTAGAAAAGCAAATCTAAAATATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594
Db 4226 TTAGAAAAGCAAATCTAAAATATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 4167
Qy 1595 TAATGACTAGAGAAGAAGATGAGATTGTCATGAATTAAGGAACGAATATTGGATA 1654
Db 4166 TAATGACTAGAGAAGAAGATGAGATTGTCATGAATTAAGGAACGAATATTGGATA 4107
Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTTGGTGTGTCAGACTGATG 1714
Db 4106 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTTGGTGTGTCAGACTGATG 4047
Qy 1715 GGCCCTATTCGGATATTGAGATGATGTGTCTATGTCTCAACAGAGAAAGCAGAGTTTCAGCC 1774
Db 4046 GGCCCTATTCGGATATTGAGATGATGTGTCTATGTCTCAACAGAGAAAGCAGAGTTTCAGCC 3987
Qy 1775 ATGAATGGACAAACGGGTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 1834
Db 3986 ATGAATGGACAAACGGGTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 3927
Qy 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCGCGCTTACACATGGTGGTCAATTTTCTCTA 1894

Db 3926 TAGATTATGCATCTCAGGTGGAATCAGATTGGCGCGCTTACACATGGTCAATTTTCTCTA 3867
Qy 1895 TTTTGGCGGATTTATGATTCAGGTGGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db 3866 TTTTGGCGGATTTATGATTCAGGTGGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 3807
Qy 1955 TAGAAGCCCAAAACGTTCCAGATGCGATTGTCGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014
Db 3806 TAGAAGCCCAAAACGTTCCAGATGCGATTGTCGCCCTTATCGTAGAAGAGCTGTTTGAAT 3747
Qy 2015 ATGCAGGCAAAATGCGGTAATTAATTCGTGTGAAGGACCGCAACAATTTCTACCATCCTTGA 2074
Db 3746 ATGCAGGCAAAATGCGGTAATTAATTCGTGTGAAGGACCGCAACAATTTCTACCATCCTTGA 3687
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGTTGTCATCATCGCATCTGTTATA 2134
Db 3686 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGTTGTCATCATCGCATCTGTTATA 3627
Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194
Db 3626 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 3567
Qy 2195 ACCATCTGTGCCAGTTTCGTAATGTCTGTCAACTTTTCCGACTCTGAGAAAATCTTCCTGGAAT 2254
Db 3566 ACCATCTGTGCCAGTTTCGTAATGTCTGTCAACTTTTCCGACTCTGAGAAAATCTTCCTGGAAT 3507
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACAGATATATAGTGG 2314
Db 3506 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACAGATATATAGTGG 3447
Qy 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341
Db 3446 ATGTGTCAAAACGCATACCATTTTGAA 3420

Search completed: February 12, 2006, 03:10:48
Job time : 1633 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2006, 02:43:45 ; Search time 366 Seconds

(without alignments)
7753.446 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 3156

Sequence: 1 ggatccacgaggttagcagc.....atcggtatgatgtcgaattc 3156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6240305 seqs, 449581930 residues

Total number of hits satisfying chosen parameters: 12480610

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*

- 1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
- 11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
- 12: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Ptsd. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
c	1	762.4	24.2	4775	7	US-10-401-386B-62
c	2	286.6	9.1	3348	7	US-10-793-626-3541
c	3	286.6	9.1	3830	7	US-10-793-626-3934
c	4	286.6	9.1	3876	7	US-10-793-626-3895
	5	285	9.0	999	7	US-10-793-626-539
	6	285	9.0	999	7	US-10-793-626-2561
	7	172	5.4	1470	11	US-11-000-688-675
	8	114.6	3.6	3619	7	US-10-793-626-3335
	9	113	3.6	1074	7	US-10-467-657-6969
	10	107.8	3.4	1146	11	US-11-098-686-9321
	11	107.8	3.4	1457619	11	US-11-098-686-8739
	12	106.4	3.4	37507	7	US-10-522-037-2
	13	100.6	3.2	3542	7	US-10-793-626-4459
	14	95	3.0	1125	11	US-11-055-822-643
	15	84	2.7	1970	11	US-11-136-527-2288
	16	83	2.6	1134	7	US-10-793-626-2925
	17	83	2.5	2955	7	US-10-793-626-3531
	18	79.8	2.5	969	11	US-11-098-686-9320
c	19	79.8	2.5	1457619	11	US-11-098-686-8739
	20	71.8	2.3	1049	11	US-11-136-527-174879
	21	71.8	2.3	1049	11	US-11-136-527-174880

Sequence 174881,	22	71.8	2.3	1049	11	US-11-136-527-174881
Sequence 174905,	23	71.8	2.3	1049	11	US-11-136-527-174905
Sequence 174906,	24	71.8	2.3	1049	11	US-11-136-527-174906
Sequence 174907,	25	71.8	2.3	1049	11	US-11-136-527-174907
Sequence 174908,	26	71.8	2.3	1049	11	US-11-136-527-174908
Sequence 174909,	27	71.8	2.3	1049	11	US-11-136-527-174909
Sequence 174910,	28	71.8	2.3	1049	11	US-11-136-527-174910
Sequence 174911,	29	71.8	2.3	1049	11	US-11-136-527-174911
Sequence 174912,	30	71.8	2.3	1049	11	US-11-136-527-174912
Sequence 174913,	31	71.8	2.3	1049	11	US-11-136-527-174913
Sequence 174914,	32	71.8	2.3	1049	11	US-11-136-527-174914
Sequence 174915,	33	71.8	2.3	1049	11	US-11-136-527-174915
Sequence 174916,	34	71.8	2.3	1049	11	US-11-136-527-174916
Sequence 174917,	35	71.8	2.3	1049	11	US-11-136-527-174917
Sequence 174918,	36	71.8	2.3	1049	11	US-11-136-527-174918
Sequence 174919,	37	71.8	2.3	1049	11	US-11-136-527-174919
Sequence 3592, Ap	38	71.8	2.3	1049	11	US-11-128-061-3592
Sequence 3593, Ap	39	71.8	2.3	1049	11	US-11-128-061-3593
Sequence 3594, Ap	40	71.8	2.3	1049	11	US-11-128-061-3594
Sequence 3592, Ap	41	71.8	2.3	1049	11	US-11-128-049-3592
Sequence 3593, Ap	42	71.8	2.3	1049	11	US-11-128-049-3593
Sequence 3594, Ap	43	71.8	2.3	1049	11	US-11-128-049-3594
Sequence 7236, Ap	44	63.2	2.0	333	11	US-11-128-061-7236
Sequence 7236, Ap	45	63.2	2.0	333	11	US-11-128-049-7236

ALIGNMENTS

RESULT 1
US-10-401-386B-62/c
; Sequence 62, Application US/10401386B
; Publication No. US20050261213A1
; GENERAL INFORMATION:
; APPLICANT: Patrick Branigan
; APPLICANT: Theresa J Goletz
; APPLICANT: David M Knight
; APPLICANT: Stephen G McCarthy
; APPLICANT: Bernard J Scallion
; APPLICANT: Linda A Snyder
; TITLE OF INVENTION: Nucleic Acid Compositions and Methods
; FILE REFERENCE: for Use
; FILE REFERENCE: CEN 310CIP
; CURRENT APPLICATION NUMBER: US/10/401.386B
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: 10/247,203
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 60/328,371
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 4775
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1177)
; OTHER INFORMATION: Kanamycin resistance
; FEATURE:
; NAME/KEY: rep origin
; LOCATION: (1178)...(2129)
; FEATURE:
; NAME/KEY: enhancer
; LOCATION: (2371)...(2663)
; OTHER INFORMATION: Rous Sarcoma Virus enhancer
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (2766)...(3508)
; OTHER INFORMATION: HCMV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3509)...(4533)

OTHER INFORMATION: PSA CDNA

FEATURE:
NAME/KEY: polyA site
LOCATION: (4580)...(4775)
OTHER INFORMATION: SV40
US-10-401-386B-62

```
Query Match      24.2%; Score 762.4; DB 7; Length 4775;
Best Local Similarity 96.8%; Pred. No. 7.8e-182;
Matches 778; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1564 AAGGGAATGAGAATAGTAATGGAACCAATAATAATGACTAGAGAAGAAAGAAATGAAGATT 1623
Db      |||
Qy 1074 AATCAAGGGGTGTTATGAACGGACCAATAATAATGACTAGAGAAGAAAGAAATGAAGATT 1015
Db      |||
Qy 1624 GTTCATGAATAATAGGAACGAATATTGGATAAATAATGGGGATGATGTTAAGGCTATTGGT 1683
Db      |||
Qy 1014 GTTCATGAATAATAGGAACGAATATTGGATAAATAATGGGGATGATGTTAAGGCTATTGGT 955
Db      |||
Qy 1684 GTTTATGGCTCTCTTGGTGGTTCAGACTGATCGGCCCTATTTCGGATATTGAGATGATGTT 1743
Db      |||
Qy 954 GTTTATGGCTCTCTTGGTGGTTCAGACTGATCGGCCCTATTTCGGATATTGAGATGATGTT 895
Db      |||
Qy 1744 GTCATGTCAACAGAGGAACGAGGTTTCAGCCATGAATGGAACACCGGTGAGTGGAAAGGTG 1803
Db      |||
Qy 894 GTCATGTCAACAGAGGAACGAGGTTTCAGCCATGAATGGAACACCGGTGAGTGGAAAGGTG 835
Db      |||
Qy 1804 GAAGTGAATTTTGATAGGCAAGAGATTCTACTAGATTATGATCTCAGGTGGAATTCAGAT 1863
Db      |||
Qy 834 GAAGTGAATTTTGATAGGCAAGAGATTCTACTAGATTATGATCTCAGGTGGAATTCAGAT 775
Db      |||
Qy 1864 TGGCGCTTACACATGGTCAATTTTCTATTTTCCCGATTATGATTCAGGTGGATAC 1923
Db      |||
Qy 774 TGGCGCTTACACATGGTCAATTTTCTATTTTCCCGATTATGATTCAGGTGGATAC 715
Db      |||
Qy 1924 TTAGAGAAGTGTATCAAACTGCTAAATCGGTAGAAAGCCCAAGCTTCCAGATGCCATT 1983
Db      |||
Qy 714 TTAGAGAAGTGTATCAAACTGCTAAATCGGTAGAAAGCCCAAGCTTCCAGATGCCATT 655
Db      |||
Qy 1984 TGTGCCCTTATGTAGAAAGAGCTGTTTGAATATGCAAGCAAAATGCGTAATATTTCGTGTG 2043
Db      |||
Qy 654 TGTGCCCTTATGTAGAAAGAGCTGTTTGAATATGCAAGCAAAATGCGTAATATTTCGTGTG 595
Db      |||
Qy 2044 CAAGGACCGACAAATTTCTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCCATG 2103
Db      |||
Qy 594 CAAGGACCGACAAATTTCTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCCATG 535
Db      |||
Qy 2104 TTGATTGGTCTGCATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAGCA 2163
Db      |||
Qy 534 TTGATTGGTCTGCATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAGCA 475
Db      |||
Qy 2164 GTTAAGCAATCAGATCTTCTTCAGGTTATGACCAATCTGTGCCAGTTCGTAATGTCTGGT 2223
Db      |||
Qy 474 GTTAAGCAATCAGATCTTCTTCAGGTTATGACCAATCTGTGCCAGTTCGTAATGTCTGGT 415
Db      |||
Qy 2224 CAATTTCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAG 2283
Db      |||
Qy 414 CAATTTCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAG 355
Db      |||
Qy 2284 GAGTGGACAGAACGACACGGATATATAGTGGATGTGTCAAAACGCATACCAATTTTGAATT 2343
Db      |||
Qy 354 GAGTGGACAGAACGACACGGATATATAGTGGATGTGTCAAAACGCATACCAATTTTGAACG 295
Db      |||
Qy 2344 CGAAAGCGCCGATTGAGTCTTACC 2367
Db      |||
Qy 294 CAGAATTGGTTAATTGGTTGTAAC 271
Db      |||
```

RESULT 2

US-10-793-626-3541/c
Sequence 3541, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:

APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3541
LENGTH: 3348
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3541

```
Query Match      9.1%; Score 286.6; DB 7; Length 3348;
Best Local Similarity 61.8%; Pred. No. 4.2e-62;
Matches 473; Conservative 0; Mismatches 289; Indels 3; Gaps 1;

Qy 2371 TGGTGAATAAGGAACACCTGCTTGAAGCGCGAAGCGGCGCACGATCTGAATATCGGCA 2430
Db      |||
Qy 2492 TTGTTGAACAGGACCAAAATTAAGAAGCGCTCAAGTGGCAACTGAAAAATCAAAATCGGTA 2433
Db      |||
Qy 2431 CATATTGTCGTGGCAAGCGGAGAGTCCGTCTAACAGAGAGTGGATCAGTCTCGTAG 2490
Db      |||
Qy 2432 CATACTGATTGTTATGAGTGGTAGAGTCTTAGTAAACAGAGAGTGCATCATATTGGCG 2373
Db      |||
Qy 2491 ATGCGGTTCAGGAAATTAAGAGACGTA---TGGACTGAAGATTGTGCATGCTCTGGAC 2547
Db      |||
Qy 2372 AAACAGTAGAAGATATTAAAGAGATACACCCACNACTAAAGATTGTGCGTGTAGGAT 2313
Db      |||
Qy 2548 TGTGAAGCCAGACGAGCGGCTCAAAGATCAGAGATAGACCGCTATATATCATATA 2607
Db      |||
Qy 2312 TAAAGAAAGAAAGAACAGGCTAAAGAAATTAAGAGGCTGCTGGTGCATCGTTATAATCATA 2253
Db      |||
Qy 2608 ATTTGAATAGCTACAGAGAAACCAATTCAAACATCAACACTCACATACATACATGACA 2667
Db      |||
Qy 2252 ATTAAATACGAGTGAGCGTTATCACGATGAAGTAGTAATACACATACATATGAGGATA 2193
Db      |||
Qy 2668 GAGTCAATACGTTGAAATCGCAAAAGAAATCGGGGCTGTCTCCGTGTTTCAGGCGCCATT 2727
Db      |||
Qy 2192 GAGTGAATACGTTGAAATGATGAAGATAATAATAATTTCTCTTGTTCAGTGTGATAT 2133
Db      |||
Qy 2728 TCGGGAATGAAGGAGACGAAACAGAGATGTCATGACATCGCCAAAGCTTGAAGGCTCTTG 2787
Db      |||
Qy 2132 GTGGTATGGGAGAGTCCGAATGAGGACATATTATGATATGGCATTTTGCTTTAAGAGCCATCG 2073
Db      |||
Qy 2788 ACGGGATTCATTCCTGTCGAATTTTTCGATGCAATTTGATGCGCAAGCTTGAAGGCG 2847
Db      |||
Qy 2072 ATGCTGATAGCATTCCTATTAAATTTTTCATCTCTTAAAGGAACATAATTTTGGTGGAT 2013
Db      |||
Qy 2848 TCACGAATTTAAACCCCTGTTATTTTAAAGTGCTGGCGCTGTTCGCTTTTATCAATC 2907
Db      |||
Qy 2012 TAGATTTATTGTCCACCAATGAATGTTTAAGATATAGCGATGTTTAGGTTATCAATC 1953
Db      |||
Qy 2908 CATCAAAAGAAATTCGCAATTTCCGAGAGAAAGAGAGTCAATCTCCGACATGTCAGGCAAT 2967
Db      |||
Qy 1952 CAAACAAAGAAATTCGAATTTGCAAGTTGCAAGTGGAGAGTAAATCTACGTTTCAATCAACCCAC 1893
Db      |||
Qy 2968 TAGGGCTTTACGGCGCAACTCCATTTTTCGAGAGACTACTTAAACACTGCCGGGCAAG 3027
Db      |||
Qy 1892 TCGCATTTGAAAGCGGCTAAATTCAAATTTTGTAGGAGATTACTTAAATTACAGCGGCTCAAC 1833
Db      |||
Qy 3028 AGGAGACGAGAGATCATAAAAATGCTGAGTGAATTTAGGCTTTGAAAGTTCAGTTCGAAG 3087
Db      |||
Qy 1832 CGATGAGGAAGATTATCGCATGATTGAAGATTTAGGCTTTGAAATCGACAGATTAAATTA 1773
Db      |||
Qy 3088 AAATGAAGGCTAGTTTAAGTGGCAAAAGCTGAAAGAAATCAATAAA 3132
Db      |||
Qy 1772 TAAATTAATAATATTTTCAATAAATGTAATATATAGTTTAAATAAA 1728
Db      |||
```


Db 2577 CAAACAAAGAAATTCGAATTCAGGTGCGACGGGAGGTAAATCTAGCTTCATTATCAACCCAC 2518
Qy 2968 TAGGGCTTTACGCGCAACTCCATTTTTCGCGAGACTACTTAAACAACCTCCCGGCAAG 3027
Db 2517 TCGCATTTAAAGCGGCTAAATTCATTTTGTGAGAGATTTACTTAATTACAGCGCGTCAAC 2458
Qy 3028 AGGAGACGGAGATCATAAATGCTGAGTGATTTAGGCTTTGAAAGTTGAATCACTCGAAG 3087
Db 2457 CGAATGAGGAAGATTATCGCATGATTGAGATTTAGGGTTTGAATCGACAGITTAATTA 2398
Qy 3088 AAATGAAGGCTAGTTTAAGTCGAAAAAGCTGAAAGAAATCAATAAA 3132
Db 2397 TAAATTAATAATTTTCAATAAATGTAATATATAGTTTAAATAAA 2353

RESULT 5

US-10-793-626-539
; Sequence 539, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 539
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-539

Query Match 9.0%; Score 285; DB 7; Length 999;
Best Local Similarity 63.6%; Pred. No. 6.5e-62;
Matches 451; Conservative 0; Mismatches 255; Indels 3; Gaps 1;
Qy 2371 TGGTGAATTAAGGAAACGCTGCTTGAAGCGCGGAAGCGGCGCAGCATCTGAATATCGCA 2430
Db 284 TTGTTGAACAGGACCAAAATTAAGAGAGCGCTCAAGTGGCAACTGAAAAATCAAAATCGTA 343
Qy 2431 CATATTGTATCGTGGCAAGCGGCGAGAGTCCGCTCTAACAGAGAAAGTGGATCGGTAG 2490
Db 344 CATACTGTATTGTTATGAGTGGTAGAGTCTTAGTAACAGAGAAAGTGCATCATATTTGCG 403
Qy 2491 ATGGGTTTCAGGAATTAAGAGAGGTA---TGGACTGAAGATTTGTGCATGCTTGGAC 2547
Db 404 AAACAGTAGAAGATTAATAAGAGATACACCCACAACCTAAAGATTTGTGCGTGTAGAT 463
Qy 2548 TGTGTAAGCAGAGCAGCGGCTCAAGATGAGGAGTAGACCGCTATATATCATATA 2607
Db 464 TAACGAAGAAGAACAGGCTAAAAAATTAAGGCTGCTGTGTCGATCGTTATATATATA 523
Qy 2608 ATTTGAATACGTCAAGAGAAACCATTCACACATCAACACCTCACATACATACATGACA 2667
Db 524 ATTTAAATACAGTGAGCGGTTATCACGATGAAGTAGTAACACATACATATAGGATA 583
Qy 2668 GAGTCAATACGGTTGAAATCCGAAAGATCCGGGGTGTCTCCGTGTTTCAGGCCCATTA 2727
Db 584 GAGTGAATACGGTTGAAATGATGAAGATAATAATATTTCTCTTTGTCAGGTGTGATAT 643
Qy 2728 TCGGGATGAAGGAGCAGAAACAGGATGTCTTGAATCGCCAAAGCTTGAAGGCTCTTG 2787
Db 644 GTGGTATGGGAGATCGAATGAGGACATTTATGATATGGCAATTTGCTTTAAGAGCCATCG 703
Qy 2788 ACGGGATTCATTCCTGTGTAATTTTTTGCATGCAATTTGATGGCAGCGCGTTAGAAGCG 2847

Db 704 ATGCTGATAGCATTCCTATTAATTTTTTACATCCTATTAAGGAACATAAATTTGGTGGAT 763
Qy 2848 TCAACGAATTAACCCGCTGTATTTGTTTAAAGTGTGCGCTGCTCCGTTTATCAATC 2907
Db 764 TAGATTTATGTCAACCAATGAATGTTTAAAGATTAAGCATTTAGGTATTAATCAATC 823
Qy 2908 CATCAAAAGAAATTCGCAATTTCCGGAGGAAGAGAGGTCAATCTCCGACATTCAGGCCAT 2967
Db 824 CAACAAAAGAAATTCGAATTCAGGTGCGAGGAGTAAATCTACGTTCAATTAACACCAC 883
Qy 2968 TAGGGCTTTAGCGCGCAAACTCCATTTTGTGCGAGACTACTTAAACAACCTCCCGGCAAG 3027
Db 884 TCGCATTTGAAAGCGGCTAAATTCATTTTGTAGGAGATTTACTTTAATTTACAGCGCGTCAAC 943
Qy 3028 AGGAGACGGAGGATCATAAATGCTGAGTGATTTAGGCTTTGAAGTTGA 3076
Db 944 CGAATGAGGAAGATTATCGCATGATTGAAGATTTAGGGTTTGAATCGA 992

RESULT 6

US-10-793-626-2561
; Sequence 2561, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2561
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2561

Query Match 9.0%; Score 285; DB 7; Length 999;
Best Local Similarity 63.6%; Pred. No. 6.5e-62;
Matches 451; Conservative 0; Mismatches 255; Indels 3; Gaps 1;
Qy 2371 TGGTGAATTAAGGAAACGCTGCTTGAAGCGCGGAAGCGGCGCAGCATCTGAATATCGCA 2430
Db 284 TTGTTGAACAGGACCAAAATTAAGAGAGCGCTCAAGTGGCAACTGAAAAATCAAAATCGTA 343
Qy 2431 CATATTGTATCGTGGCAAGCGGCGAGAGTCCGCTCTAACAGAGAAAGTGCATGCTTGGAC 2490
Db 344 CATACTGTATTGTTATGAGTGGTAGAGTCTTAGTAACAGAGAAAGTGCATCATATTTGCG 403
Qy 2491 ATGGGTTTCAGGAATTAAGAGAGGTA---TGGACTGAAGATTTGTGCATGCTTGGAC 2547
Db 404 AAACAGTAGAAGATTAATAAGAGATACACCCACAACCTAAAGATTTGTGCGTGTAGAT 463
Qy 2548 TGTGTAAGCAGAGCAGCGGCTCAAGATGAGGAGTAGACCGCTATATATCATATA 2607
Db 464 TAACGAAGAAGAACAGGCTAAAAAATTAAGGCTGCTGTGTCGATCGTTATATATATA 523
Qy 2608 ATTTGAATACGTCAAGAGAAACCATTCACACATCAACACCTCACATACATACATGACA 2667
Db 524 ATTTAAATACAGTGAGCGGTTATCACGATGAAGTAGTAACACATACATATAGGATA 583
Qy 2668 GAGTCAATACGGTTGAAATCCGAAAGATCCGGGGTGTCTCCGTGTTTCAGGCCCATTA 2727
Db 584 GAGTGAATACGGTTGAAATGATGAAGATAATAATATTTCTCTTTGTCAGGTGTGATAT 643
Qy 2728 TCGGGATGAAGGAGCAGAAACAGGATGTCTTGAATCGCCAAAGCTTGAAGGCTCTTG 2787
Db 644 GTGGTATGGGAGATCGAATGAGGACATTTATGATATGGCAATTTGCTTTAAGAGCCATCG 703

Qy	2788	ACGCGGATTCGATTCCTGTGAAATTTTTTTTGATGCAATTGATGCGACGCCGTTTGAAGCG	2847
Db	704	ATGCTGATAGCATTCCTATTAAATTTTTTACATCCTATTTAAAGGAACATAAATTTGGTGGAT	763
Qy	2848	TCACGAATTAACACCGCGCTGATTGTTTTAAAAGTGCTGGCGCTGTTCCGTTTTTATCAATC	2907
Db	764	TAGATTATTGTGCACCAATGAAATGTTTTAAGAAATTATAGCGATGTTTGGTTTAATCAATC	823
Qy	2908	CATCAAAGAAATTCGCAATTTCCGCGAGAGAGAGGTCAATCTCCGCACATTCGAGCGCAT	2967
Db	824	CAACAAAGAAATTCGAATTTGCAGTTCGACGGGAGGTAAATCTACGTTCTATCAACCCAC	883
Qy	2968	TAGGGCTTTAGCGCGCAAACTCCATTTTTTTGTCGAGACTACTTAACAACTGCCGGGGAAG	3027
Db	884	TCGCATTTGAAGCGGGCTAAATTCAAATTTTTGTGAGAGATTACTTAAATTCAGGGCGGTCAAC	943
Qy	3028	AGGAGACGGAGGATCATAAAATGCTGAGTGATTTTAGGCTTTTGAAGTTGA	3076
Db	944	CGAATGAGGAGATTTATCGCATGATCGAAGATTTAGCGTTTGAATTCGA	992

RESULT 7

```

US-11-000-688-675
: Sequence 675, Application US/11000688
: Publication No. US20050287544A1
: GENERAL INFORMATION:
: APPLICANT: BERTUCCI, Francois
: APPLICANT: HOULGAITE, Remi
: APPLICANT: BIRNSAUM, Daniel
: TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
: FILE REFERENCE: 1423-R-03
: CURRENT APPLICATION NUMBER: US/11/000,688
: CURRENT FILING DATE: 2004-12-01
: PRIOR APPLICATION NUMBER: US 60/525,987
: PRIOR FILING DATE: 2003-12-01
: NUMBER OF SEQ ID NOS: 1596
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 675
: LENGTH: 1470
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial sequences:primer
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1)..(1470)
: OTHER INFORMATION: glycine c-acetyltransferase
: OTHER INFORMATION: (2-amino-3-ketobutyrate coenzyme a ligase) (GCAT)
: OTHER INFORMATION: gene.
US-11-000-688-675

```

Query Match	5.4%	Score 172;	DB 11;	Length 1470;
Best Local Similarity	49.7%;	Prod. No. 2.1e-33;		
Matches 473;	Conservative 470;	Mismatches 9;	Gaps 1;	
Qy	158	CAAAACAATATATTAGGGCTCGCAAGCAGATAGACGTTTGATCGATGCGAGCCCAACACAGCAT	217	
Db	235	CCAACAACCTACCTGGGCTGAGCAGCCACCCCTGAGGTGATCCAGCAGGCTCTGCAGGCTC	294	
Qy	218	TGCAGCAATTTGGACAGGAAGCAGCGGTTCACGTTTAAACGACAGGCCAATTCGCTCTGGC	277	
Db	295	TGGAGGAGTTTGGAGCTGGCCCTCAGCTCGGTCGCGCTTTATCTGTGGAAACCCAGAGCATCC	354	
Qy	278	ATGAAAAGCTAGAAAAAGAAATTGCCAGCTTTTAAACTGCACAGAAGCGGCCCTGCTGTGTTTT	337	
Db	355	ACAAGATCTAGAACGAAAAATAGCCGCTTCCACCAGCGGAGGATGCCATCTCTATC	414	
Qy	338	CGAGGGTTACTTGGCCCAATGCTCGGTGTCCTTTTCATCTTTCGCAGAAAAAGAAATGTCA	397	
Db	415	CCAGCTGTATTGACGCCCAACCCCGGCCCTTCCTTTGAGGGCCCTCTCTGACCCCCAGAGACCGAG	474	
Qy	398	TTTTAAAGTGACCAAGCTCAATCATGCAAGTATGATCGAGGGCTGCCGACTTTTCTAAGGCTG	457	

Db	475	TCCTGTCGGACGAGCTGACCATGCTCTCCATCATGACGGCATCGGCTGTGCAGAGGCC	534
Qy	458	ATACAGTTGTTTATTCGGCATATTTGATATGAATGATCTTTGAAAAAACAAGCTGAATGAAACAC	517
Db	535	ACAAGTACCGCTATTCGCCACCTCGGACATGCCGACCTTAGAAGCCAAGCTGCAGGAGGCC	594
Qy	518	AGCGTTATCAGCGCGGTTTTATCGTAAACAGACGGAGTATTCAGCATGGATGGCACATCG	577
Db	595	AGAAGCATTCGGCTGCGCTCGTGGCCACTGATGGGGCTTTTCCATGGATGGCGACATCG	654
Qy	578	CCCCTCTTGATCAGATCATCTCACTTTCGAAAAAGCTATCATGCTTCGTGGTGGTTGATG	633
Db	655	CACCCCTGCAGGAGATCTGCTGCTCGCTCTAGATATGTTGTCCTGCTTCTTATGATG	714
Qy	638	ATGCCACGCAACAGAGATTTTGGCGGATTTGGGACAAAGAAACGAGTGAATATCTTGGTG	697
Db	715	AATGCCATGCCACTGGCTTCTGGGGCCACAGGACGGGGCACAGATGAGCTGCTGGGTG	774
Qy	698	TTTGTCCTCGAGATGTTGTTATCGGCACCTTAAGCAAAAG-----CTGTTGGCGCGGAAG	748
Db	775	TGATGGACCAAGTGCAACCATCATCAACTCCA CCGCTGGGGAAAGGCCCTCGGTGGAGCATCAG	834
Qy	749	GAGGTTTTGGCGGAGGATCAGCGGTCTTCATCGACTTTTGTGAACCATGCGCAGAACAT	808
Db	835	GGGGCTTACAGCAGAGGGCCCTGGGCCCTGTGTCTGCTCGGAGCGCGCCGCGGCAT	894
Qy	809	TTATCTTTTCAAAACCGCTATTTCCGCGACCGCAGCTGTGCGGCTGCTCAGAGGCTTTCAACA	868
Db	895	ACCTCTTCTCCACAGTCTGCCACCTGCTGTGTTGGCTCGCGCTCCAAAGGCCCTAGATC	954
Qy	869	TCATTGAAGCCAGCAGGGGAAAAACGACAGCTTTTATTTTCTTATATCAGCATGATCAGAA	928
Db	955	TGCTGATGGGGAGTAAACACCATTTGTCAGTCTATGGCTGCGCAAGACCCAGAGGTTCCGTA	1014
Qy	929	CCAGTCTGAAGATATGGGTTATGTGGTGAAGAGAGATCACACACCGATTATTCCTCTAG	988
Db	1015	GTAAGATGGAGCTGTGGCTTCACTATCTCGGGAGCCAAGTCAACCCCATCTGCCCTGTGA	1074
Qy	989	TCATTGGCGATGCCCATAAAAACGGTCTCTATTGTCTGAAAAAAGCTGCAGGCAAGGGAATTT	1048
Db	1075	TGCTGGGTGATGCCCGCTGGCTCTCGCATGGCGGATGATGCTCTGAAGAGGCATCT	1134
Qy	1049	ATGCTCCTCGCATTTGGCGCGCAACCGTTGGCGGGGTGAAGCCGATTCG	1100
Db	1135	TTGTGATCGGGTTTCAGCTACCCCGTGTGCTCCCAAGGCAAGGCCGCGATCCG	1186

RESULT 8

```

RESU01 8
US-10-793-626-3335
; Sequence 3335, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,358
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3335
; LENGTH: 3619
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3335

```

Query Match 3.6%; Score 114.6; DB 7; Length 3619;
Best Local Similarity 62.6%; Pred. No. 8.7e-19;

499 AACAGCTGAATGAACACACAGCGTTATCAGCGCGT---TTTATCGTAACAGACGAGTA 555
Db TCCCTTTTAAAAAATATAAAGATGATAATAGACCTAAATTTATGTTACAGAAACACTC 525
Qy TTACAGTGGATGACCAATCGCCCTCTTGTATCAGATCATCTCACTTGGCAACGGTAT 615
Db TTTGGCATGGATGAGACATAGCACCCCTTACAGATATGTTTTCACCTTCATCGCAATAT 585
Qy CATGCTTCGGTGGTGTGATGATGCCCAACAGGAGTTTGGCGGATTCGGGACAA 675
Db CAAACATTTTGTACTTAGCAGAGCTCATGCTACAGCTTACAGGTATTCAGGTATTCAGGATAT 645
Qy GG-----AACGAGTGAATACCTTTGGTGTGTTGCCGACATGTTATCGGCACCTTAAGC 729
Db GGCCTATCAACACAGTAATCTCAGTCATATACCATATAGTTATGGGTACTTTTAGT 705
Qy AAAGCTGTTGGCGGAGAGGTTTGGCGGAGATCAGGGTCTTCATCGACTTTTGG 789
Db AAAGCACTTGGATGTAGTGGCGGATACATAGCTTTGTAAGCAACTTATTAAGGATTAACCTA 765
Qy CTGAACCATGCCAACAATTTATCTTTCAACACGCTATTCGCCAGCCAGCTGTGCGGCT 849
Db ATAAATTAAGTACTGGCTTTATTTATTCACAGAGGCTCTCTATGGTTTAGGAGCA 825
Qy GCTCAGCAGGCTTTCAACATCAITGAAGCCAGCAGGGAACCAACAGCTTTTATTTTCT 909
Db GCACAAGACGATGGAATCTTATAAAGACACTAGATAACCAAGCTGCACACCTCAACTCT 885
Qy TATATCAGCATGATCAGAACAGCTCTGGAAGTAATGGGTATGTTGGTGAAGGAGATCAC 969
Db TTAAGTAAAGAAATPACGAGATGGGTGAACCAACTAGGATTTAATATAGGAACCTAGTTCT 945
Qy ACACCGATTATCTGTAGTCACTTGGCGATGCCCATAAACGGTCTTATTTGCTGAAAA 1029
Db TCACATATTTCTCTTATTTAGGCTCAGAAAGATGATCATCTCANGTTTCAAGAAA 1005
Qy CTGCAGGCGAAGGAAATTTATGCTCTCGCCATTCGGCGCCCAACCGTTGGCGCGGTGAA 1089
Db CTTTGAACACCATATCCCTTACCTCTTATATCAGACCTCCAACTGTTCTCCCAAGACG 1065

RESULT 11

US-11-098-686-8739
; Sequence 8739, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8739
; LENGTH: 1457619
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-8739

Query Match 3.4%; Score 107.8; DB 11; Length 1457619;
Best Local Similarity 47.0%; Pred. No. 4.9e-16;
Matches 457; Conservative 0; Mismatches 492; Indels 24; Gaps 3;

Qy 154 TCCTCAAAACAATTTATTAGGGCTCGCAAGCGATAGAGCTTTGTATGATGAGCGCCAAACA 213
Db TCTACAAATGACTATCTACAAATTAGCTTATCACCTCAACCTTATATGCTGCAATATG 516116
Qy 214 GCATTGCAGCAATTTGGGACAGACAGCGGTTTCAGGTTTAAACGACGAGCAATTCGGTC 273
Db ACAGGTAGTACCTATGAGTAGGATCCAGGGCTTAGACTTCTTTCGGAATAATATGAA 516176
Qy 274 TGGCATGAAAAGCTAGAAAAGAGATTGCCAGCTTTAAACTGACAGAGGCGGCTGCTG 333
Db CTATTTGAACGCTAGAAAACCTACTATTGCAACAAGATAAACAATCACTGAAACAGCTATGCTG 516236
Qy 334 TTTTCGAGCGGTTACTTTGGCCAATGTGCGTGTCTTTTCATCTCTTCCAG----- 382
Db TTTATATCTGGTTTCAAAACAAATGTAAAGTGTATTAATCAGCATTTATTAGATCAACATGTA 516296
Qy 383 -----AAAAGCAAGATGTCATTTTAAAGTACAGCTCAATCATCAAGTATGATCGAGGC 438
Db TTTAAATATGCAACCTCTTTGTTATTTTGTAAAGTTAAACCATGCAAGCTCTATCAAGCT 516356
Qy 439 TGCCGACTTTCTAAGGCTGATACAGTTGTTTATCGGCATATTGATATGAATGATCTTGAA 498
Db GTTTTATTAAGTAAAGCTGAACTTCTTCGTTATTATCACAATAATATGGAACACTATCA 516416
Qy 499 AACAGCTGAATGAACACACAGCGTTATCAGCGCGT---TTTATCGTAAACAGCGAGTA 555
Db TCCCTTTTAAAAAATATAAAGATGATAATAGACCTAAATTTATTTGTACAGAAACACTC 516476
Qy 556 TTACAGATGATGACCAATCGCCCTCTTGTATCAGATCATCTCACTTGGCAACGCTAT 615
Db TTTGGCATGATGAGACATAGCACCCCTTTACAGATAATGTTTTCACCTTTCATCGCAATAT 516536
Qy 616 CATGCTTCTGTGTGTTGATGATGCCAGCAACAGGAGTTTGGCGGATTCGGGACAA 675
Db GGCTATCAACAACAGTAAATCTCAGTCAATATACCATATAGTTATGGGTACTTTTAGT 516596
Qy 730 AAAGCTGTTGGCGGAGAGGTTTGGCGGAGGATCAGCGGTCTTCATCGACTTTTGG 789
Db AAAGCACTTGGATGTAGTGGCGGATACATAGCTTTGTAAGCAACTTATTAAGGATTAACCTA 516716
Qy 790 CTGAACCATGCCAGAACATTTATCTTTCAACCGCTATTCGCCAGCCAGCTGTGCGGCT 849
Db ATAAATAAGCTACTGGCTTTATTTATTCACAGAGGCTCTCTATGGTTTAGGAGCA 516776
Qy 850 GCTCAGGAGCTTTCAACATCATTTGAAGCCAGCAGGGAACCAACAGACTTTTATTTTCT 909
Db TATATCAGCATGATCAGAACCGTCTGAAGAAATATGGGTATGTTGGTGAAGGAGATCAC 969
Qy 910 TTAAGTAAAGAAATTAACAGATGGGTGTAAACCAACTAGGATTTAATATAGGAACCTAGTTCT 516836
Qy 970 ACACCGATTTTCTCTGTAGTCACTTGGCGATGCCCAATAAAGCGTCTTATTTGCTGAAAA 1029
Db TCACATATTTATCTCTTTATTTAGGCTCAGAAAGATGATCACTCAATGTTTCAAGAAA 516956
Qy 1030 CTGCAGGCGCAAGGGAATTTATGCTCTCGCATTCGGCGCCCAACCGTTGGCGCGGTGAA 1089
Db CTTTGAACACCATATCTTACTCTTATATCAGACCTCCAACTGTTCTCTCCAAAGACG 517016
Qy 1090 AGCCGATTCGAA 1102
Db TCTCGTATTCGTA 517029

RESULT 12
US-10-522-037-2
; Sequence 2, Application US/10522037
; Publication No. US20050282166A1

D _b	1143	TAGGACTATATGGAAACCGGGGTGCAGGATCGGGAGCGTGATGGAAATTATGCCAAGC	1202
Q _y	710	TTGTATTAC-----GGCACCTTAGCAAGCTGTTCGGCGCGGAAGGAGGTTTTGCGGCAG	763
D _b	1203	TTGACATCATCTCTGGAACTCTTGGCAAGGCGCTTTGGCTGTGTGCGTGGGTATATCGCCA	1262
Q _y	764	GATCAGCGGTCTTCATCGACTTTTTGCTGAACCATGCCAGAACATTTATCTTTCAAACCG	823
D _b	1263	GCAC TCGGAGCTTGGTGGACATGGTGGCGCTCTACGCTGCAGGCTTCATCTTTACCACCTT	1322
Q _y	824	CTATTTCGCCACGCCAGCTGTGCGGCTGCTCAcAGGCTTTT	863
D _b	1323	CAC TGCCCTCCCATGGTGTCTCTCGAGCTCTAGAGTCTGT	1362

Search completed: February 12, 2006, 03:17:01
Job time : 371 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frameplus_n2p_modely

Run on: February 10, 2006, 02:17:17 ; Search time 92.7 Seconds

(without alignments)
2991.757 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggttacgacg.....atcggtatgatgcgaattc 3156

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p_model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10681086/runat_09022006_140514_11701/app_query.fasta.1
-DB=A_Geneseq -QMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10681086 @CGN 1.1.348 @runat_09022006_140514_11701 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT_DSPBLOCK=100 -LONGLOG_DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

	A_Geneseq_21.*
1:	Geneseq1980s.*
2:	Geneseq1990s.*
3:	Geneseq2000s.*
4:	Geneseq2001s.*
5:	Geneseq2002s.*
6:	Geneseq2003as.*
7:	Geneseq2003bs.*
8:	Geneseq2004s.*
9:	Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1836	33.3	389	8 ADS44825	Ads44825 Bacterial
2	1333	24.2	253	2 AAW31408	Aw31408 Ant (4')-I
3	1333	24.2	253	2 AAW31402	Aw31402 Ant (4')-I
4	1333	24.2	253	4 AAB62515	Aab62515 S. aureus
5	1333	24.2	253	4 AAB29794	Aab29794 S. aureus
6	1318	23.9	253	4 AAB62512	Aab62512 S. aureus
7	1267	23.0	253	4 AAB62513	Aab62513 Mutant ka
8	1222	22.1	253	4 AAB62514	Aab62514 Mutant ka
9	971	17.6	331	1 AAP70328	Aap70328 Biotin-ty

10	957	17.3	337	1	AAP81191	Aap81191 Biotin ty
11	942	17.1	387	2	AAW30520	Aw30520 Kurthia s
12	932	16.9	393	1	AAP81194	Aap81194 Biotin bi
13	851	15.4	338	2	AAW30521	Aw30521 Kurthia s
14	843	15.3	392	7	ADF41704	Adf41704 Bacillus
15	843	15.3	392	8	ADS44705	Ads44705 Bacterial
16	828	15.0	321	6	ABU43285	Abu43285 Protein e
17	828	15.0	327	5	ABP38626	Abp38626 Staphyloc
18	828	15.0	337	8	ADS05126	Ads05126 Staphyloc
19	828	15.0	332	4	AAG81723	Aag81723 S. epider
20	828	15.0	332	4	AAG82734	Aag82734 S. epider
21	809	14.7	332	6	ABU17909	Abu17909 Protein e
22	793	14.4	318	4	AAU34288	Aau34288 Staphyloc
23	793	14.4	335	4	AAU37532	Aau37532 Staphyloc
24	793	14.4	335	4	AAU37187	Aau37187 Staphyloc
25	793	14.4	335	6	ABU16098	Abu16098 Protein e
26	793	14.4	335	6	ABM73372	Abm73372 Staphyloc
27	778	14.1	321	6	ABU43752	Abu43752 Protein e
28	767	13.9	371	8	ADS44565	Ads44565 Bacterial
29	712	12.9	392	8	ADS24461	Ads24461 Bacterial
30	712	12.9	395	8	ADS28438	Ads28438 Bacterial
31	711	12.9	396	6	ABU18160	Abu18160 Protein e
32	710	12.9	388	8	ADS27616	Ads27616 Bacterial
33	705	12.8	395	4	AAU36848	Aau36848 Staphyloc
34	704	12.8	385	4	AAU34079	Aau34079 Staphyloc
35	704	12.8	396	6	ABU30082	Abu30082 Protein e
36	704	12.8	414	7	ADC96144	Adc96144 E. faeciu
37	700	12.7	393	8	ADS29218	Ads29218 Bacterial
38	699	12.7	324	6	ABU17623	Abu17623 Protein e
39	699	12.7	395	6	ABM72451	Abm72451 Staphyloc
40	699	12.7	398	6	ABU43750	Abu43750 Protein e
41	699	12.7	400	8	ADS41778	Ads41778 Bacterial
42	698.5	12.7	372	8	ADN18405	Adn18405 Bacterial
43	695.5	12.6	400	6	ADB08430	Adb08430 Alloioococ
44	692	12.5	398	4	AAB96462	Aab96462 Putative
45	692	12.5	398	8	ADS43254	Ads43254 Bacterial

ALIGNMENTS

RESULT 1

ADS44825
ID ADS44825 standard; protein; 389 AA.

XX ADS44825;

XX AC

DT 02-DEC-2004 (first entry)

XX DE

DE Bacterial polypeptide #23255.

XX KW

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX OS

XX Bacteria.

XX PN

PN US2003233675-A1.

XX PD

PD 18-DEC-2003.

XX PF

PF 20-FEB-2003; 2003US-00369493.

XX PR

PR 21-FEB-2002; 2002US-0360039P.

XX PA

PA (CAOV/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.


```
XX SQ Sequence 253 AA;
Alignment Scores:
Pred. No.: 1.08e-131 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x AAB62515 (1-253)
QY 1579 GTGAATGACCAATAATAATGACTAGACAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAAGCAGAGTTTCAGCCATGAATGGAACACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGTAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheAsp 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTCGCCGATTTATGATTCAGGTGATATAGAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyTyrLeuGluLysValTyr 120
QY 1939 CAACTGCTTAATCGGTAGAACCCCAACGTTTCCACGATGCGATTGTGTCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTTGAATATGACGGCAAAATGGCGTAATATTTCGTGTCAAGGACCGACAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTCACGTGTACAGTAGCAATGCGAGTGCCATGTTGATTCGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCCTCCAGGTTATGACCATCTGTGCCAGTTTCGTAATGTCTGTGTCACACTTCCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTTCGGAATCCCTAGAGAAATTTCTGGAATGGGATTCAGGAATGACAGCAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 4
ID AAB62515
XX AAB62515 standard; protein; 253 AA.
AC AAB62515;
XX 23-JUL-2001 (first entry)
XX
```

```
DE XX S. aureus wild-type kanamycin nucleotidyltransferase (KNT) .
KW XX Kanamycin nucleotidyltransferase; KNT; thermostability; mutation; mutant;
KW marker; thermophilic; Thermus thermophilus.
XX Staphylococcus aureus.
XX OS EP1097990-A2.
XX PN 09-MAY-2001.
XX PD 27-OCT-2000; 2000EP-00309475.
XX PF 29-OCT-1999; 99JP-00309616.
XX PR (RIKE ) RIKEN KK.
XX PA Yokoyama S, Hoseki J, Yano T, Koyama Y, Kuramitsu S;
PI Kagamiyama H;
XX WPI; 2001-337112/36.
XX DR N-FSDB; AAF83650.
XX Novel mutant kanamycin nucleotidyltransferase useful for screening
PT thermophilic bacteria such as Thermus thermophilus, has point mutations
PT at specific positions and improved thermostability.
XX Disclosure; Page 22-24; 30pp; English.
XX The invention relates to mutant kanamycin nucleotidyltransferases (KNT)
CC having improved thermostability. The mutant has one or more point
CC mutations selected from Met57Leu, Ala62Val, Ser94Pro, Ser203Pro,
CC Asp205Val, His207Gln, Ser220Pro, Ile234Val and Thr238Ala as against a
CC WT-KNT (AAB62515). A KNT gene encoding the mutant can be used as a
CC selective marker for thermophilic bacteria. The marker can be used for
CC screening thermophilic bacteria, preferably Thermus thermophilus. The
CC present sequence represents a S. aureus wild-type KNT protein. The WT-KNT
CC protein from which the mutants are constructed differs from this protein
CC at position 80 and position 130
XX SQ Sequence 253 AA;
Alignment Scores:
Pred. No.: 1.08e-131 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x AAB62515 (1-253)
QY 1579 GTGAATGACCAATAATAATGACTAGACAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAAGCAGAGTTTCAGCCATGAATGGAACACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGTAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheAsp 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTCGCCGATTTATGATTCAGGTGATATAGAAAGTGTAT 1938
```

Db 101 GlyGlnPheSerIleLeuProIleTyAspSerGlyGlyTyLeuGluYsValTy 120
Qy 1939 CAAACTCTAAATCGGTAGAGCCAAACGTTCCACGATCGCATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaYsSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTGAATATGAGGCAAAATGCGTAATATTCGTGTGCAAGGACCGACACA 2058
Db 141 GluGluLeuPheGluTyAlaGlyLeuTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTCTACCATCTGCTAGCTAGCTAGCAATGCGAGTGCCTGATGTTGATGTTGCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGCATCTGTATACGAGCGGCTCGTCTTACTCAAGCAGTAAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyThrThrSerAlaSerValLeuThrGluAlaValYsGlnSerAsp 200
Qy 2179 CTTCTCTCAGGTTATGACCATCTGTGCGAGTTGCTGATGCTGCTCAACTTTCGCACTCT 2238
Db 201 LeuProSerGlyTyAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GADAACTTCTGGAATCGCTAGAGAAATTTCTGGAATCGGATTCAGAGTGCAGACGAACGA 2298
Db 221 GluYsLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy 2299 CACGGATATATGATGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyIleValAspValSerLysArgIleProPhe 253

RESULT 5
AAB29794
ID AAB29794 standard; protein; 253 AA.
XX AAB29794;
AC AAB29794;
XX
DT 09-MAR-2001 (first entry)
XX
DE S. aureus adenyl 4'-nucleotidyltransferase type Ia (ant(4')-Ia).
XX
KW Composite nucleic acid molecule synthesis; ligation; bridge oligomer;
KW bridging oligonucleotide; bridge oligomer; heat-stable ligase;
KW chimeric kanamycin resistance gene; aph(3')-Ia promoter; terminator;
KW aminoglycoside 3'-phosphotransferase type Ia; modified ant(4')-Ia;
KW adenyl 4'-nucleotidyltransferase; gentamicin-sensitive;
KW selectable marker; clinical expression vector.
XX
OS Staphylococcus aureus.
XX
XX US6143527-A.
XX
XX 07-NOV-2000.
XX
XX 06-MAY-1997; 97US-00852268.
XX
XX 06-MAY-1996; 96US-00642045.
XX
XX (AMHP) AMERICAN HOME PROD CORP.
XX
XX Samuel M, Satishchandran C, Pachuk C;
XX
XX WPI; 2001-006434/01.
XX
XX N-PSDB; AAC81670.
XX
XX Producing a composite nucleic acid molecule for site specific ligation of
XX digested DNA fragments, involves ligation of digested double-stranded DNA
XX molecules by DNA ligases and bridging oligonucleotides.
XX
XX Example 1; Col 25-28; 4lpp; English.
XX
XX The invention relates to a novel method of synthesizing a composite
XX nucleic acid molecule. The method involves denaturing two separate double

CC stranded nucleic acids (N1 and N2), and hybridising them to a bridging
CC oligonucleotide comprising, in the 5' to 3' direction, a 10-40 nucleotide
CC sequence complementary to the 5' end of the first strand of N1 and a 10-
CC 40 nucleotide sequence complementary to the 3' end of the first strand of
CC N2. The N1 first strand 5' end and the N2 first strand 3' end are thus
CC oriented adjacent to each other, and a heat-stable ligase is then used to
CC ligate the N1 and N2 first strands to form the first strand of the
CC composite nucleic acid. The composite first strand is then denatured from
CC the bridging oligonucleotide, and allowed to reanneal to the
CC complementary strands of N1 and N2. The complementary strands are then
CC ligated to form a composite complementary strand. The invention also
CC encompasses the ligation of a number of nucleic acid molecules by the
CC method, so that the mixture comprises a number of nucleic acid molecules
CC and a number of bridging oligonucleotides. The method is useful for the
CC site-specific ligation of DNA fragments generated by PCR, restriction
CC enzyme digestion, DNase digestion, chemical cleavage, and enzymatic or
CC chemical synthesis. Large nucleic acid molecules are generated from
CC smaller ones without the need to treat the ends to ensure specific
CC binding of two molecules in correct orientation. The bridge oligomer
CC template ensures correct orientation and specific ligation. Ligation of
CC multiple components can take place without the need for intervening
CC steps, using a heat-stable ligase and the necessary starting materials.
CC In an exemplification of the invention, a chimeric kanamycin resistance
CC gene was synthesised using the method of the invention. The chimeric
CC kanamycin gene comprises the promoter and terminator of the Escherichia
CC coli kanamycin resistance gene aph(3')-Ia (which encodes aminoglycoside
CC 3'-phosphotransferase type Ia) and the open reading frame of the
CC Staphylococcus aureus kanamycin resistance gene ant(4')-Ia (which encodes
CC adenyl 4'-nucleotidyltransferase type Ia). The ant(4')-Ia open reading
CC frame was also modified such that the wild-type gram-positive GTG
CC translation initiation codon was altered to the gram-negative preferred
CC ATG initiation codon, and a nearby out-of-frame ATG codon was altered to
CC ACG. Neither of these alterations alter the amino acid sequence of the
CC protein as expressed in S. aureus. The chimeric kanamycin resistance gene
CC has the activity of the S. aureus ant(4')-Ia protein (including being
CC sensitive to gentamicin) but is able to be expressed in E. coli. The
CC narrower range of activity of the chimeric kanamycin resistance gene
CC against aminoglycoside antibiotics makes it particularly useful as a
CC selectable marker in the production of vectors for clinical use in
CC humans. The present sequence represents the amino acid sequence of the
CC Staphylococcus aureus ant(4')-Ia kanamycin resistance protein, which is
CC encoded in an Escherichia coli-expressible form by the chimeric kanamycin
CC resistance gene produced using the method of the invention
XX
SQ Sequence 253 AA;

Alignment Scores:
Pred. No.: 1.08e-131 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x AAB29794 (1-253)

Qy 1579 GTCAATGACCAATAATAATGACTAGAGAGAAAGAAATGATGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGAATATGGATAAATATGGGATGATGTTTAAGGCTATTGGTGTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyArgAspValLysAlaIleGlyValTyGlySerLeu 40
Qy 1699 GGTGCTCAGATCTGATGGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACACAG 1758
Db 41 GlyArgGlnThrAspGlyProTySerAspIleGluMetMetCysValMetSerThrGlu 60
Qy 1759 GAAGCAGAGTTTCAGCCATGAATCGAACCCGGTGGTGAAGGTGAAGTGAATTTTGTAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp 80
Qy 1819 AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTTGCCCTTACACAT 1878

```
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTyrProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTTCCGATTTATGATTCAGGTGGTACTTAGAGAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAAACTGCTAAATCGGTAGAACCCCAACGTTTCCACGATGCGATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTGTAATATGCGAGGCAAAATGGCGTAAATATTCGTGTCGAAGGCCGACAAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAenIleAsgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGACTGACAGTACAGTACGATGCGAGTGGCATGTTGATGGTGGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGAGTAAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
Qy 2179 CTTCCTTCAGGTATGACATCTGTGCGAGTTCGTAATGTCGTGTCGTAATCTTCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTTCGGAATCGCTAGAGATTTTCGGAATGCGATTCAGGAGTGCACAGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGlnIleThrGluArg 240
Qy 2299 CACGATATAGTGGATGTGTCAAAACGCGATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 6
AAB62512
ID AAB62512 standard; protein; 253 AA.
XX
AC AAB62512;
XX
DT 23-JUL-2001 (first entry)
XX
DE S. aureus kanamycin nucleotidyltransferase WT*KNT protein.
XX
KW Kanamycin nucleotidyltransferase; KNT; thermostability; mutation; mutant;
KW marker; thermophilic; thermus thermophilus.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 80
FT /label= D80Y
FT /note= "wild-type Asp is replaced with Tyr"
FT Misc-difference 130
FT FT T130K
FT /label= T130K
FT /note= "wild-type Thr is replaced with Lys"
XX
FN EP1097990-A2.
XX
PD 09-MAY-2001.
XX
PF 27-OCT-2000; 2000EP-00309475.
XX
PR 29-OCT-1999; 99JP-00309616.
XX
PA (RIKE ) RIKEN KK.
XX
PI Yokoyama S, Hoseki J, Yano T, Koyama Y, Kuramitsu S;
PI Kagamiyama H;
XX
DR WPI; 2001-337112/36.
XX
```

```
PT Novel mutant kanamycin nucleotidyltransferase useful for screening
PT thermophilic bacteria such as Thermus thermophilus, has point mutations
XX at specific positions and improved thermostability.
PS Claim 1; Page 12-13; 30pp; English.
XX
CC The invention relates to mutant kanamycin nucleotidyltransferases (KNT)
CC having improved thermostability. The mutant has one or more point
CC mutations selected from Met57Leu, Ala62Val, Ser94Pro, Ser203Pro,
CC Asp206Val, His207Gln, Ser220Pro, Ile234Val and Thr238Ala as against a
CC WT*KNT (AAB62512). A KNT gene encoding the mutant can be used as a
CC selective marker for thermophilic bacteria. The marker can be used for
CC screening thermophilic bacteria, preferably Thermus thermophilus. The
CC present sequence represents the S. aureus WT*KNT protein from which the
CC mutants are constructed. This differs from the wild-type KNT protein
CC (AAB62515) at position 80 and position 130
XX
SQ Sequence 253 AA;

Alignment Scores:
Pred. No.: 4.26e-130 Length: 253
Score: 1318.00 Matches: 250
Percent Similarity: 99.2% Conservative: 1
Best Local Similarity: 98.8% Mismatches: 2
Query Match: 23.9% Indels: 0
DB: 4

US-10-681-086-1 (1-3156) x AAB62512 (1-253)
Qy 1579 GTGAATCGACAATAAATAATGACTAGAGAGAAAGATGAAGATGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGAATATGGATAAATATGGGATGATGTTAAGCGTATTGGTATTGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTCTGTCAGACTGATGGGCCCTTATTCGATATTCAGATGATGTTGTCATGTCACACAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60
Qy 1759 GAACGAGATTTCAGCCCATGAATGACCAACCGGTGAGTGGAGGTGGAATGGAATTTGAT 1818
Db 61 GluAlaGluPheSerHisGluTyrThrThrGlyGluTyrLysValGluValAsnPheTyr 80
Qy 1819 AGCGAAGAGATTCTACTAGATTATGTCATCTCAGGTGGAATCAGATTCGCCCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTyrProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGATACTTAGAGAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAAACTGTAATCGGTAGAACCCCAACGTTCCACATGCGATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTGTAATATGCGAGGCAAAATGGCGTAAATATTCGTGTCGAAGGCCGACAAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAenIleAsgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGACTGACAGTACAGTACGATGCGAGTGGCATGTTGATGGTGGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGAGTAAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
Qy 2179 CTTCCTTCAGGTATGACATCTGTGCGAGTTCGTAATGTCGTGTCGTAATCTTCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
```


QY 2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTCGACAGAACGA 2298
 DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
 QY 2299 CACGGATATAGTGGATGTCGAAACGATACCATTT 2337
 DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 7
 AAB62513
 ID AAB62513 standard; protein; 253 AA.
 AC AAB62513;
 XX 23-JUL-2001 (first entry)
 XX Mutant kanamycin nucleotidyltransferase KT3-11.
 XX Kanamycin nucleotidyltransferase; KNT; thermostability; mutation; mutant;
 KW marker; thermophilic; Thermus thermophilus.
 XX Staphylococcus aureus.
 OS Synthetic.
 XX EP1097990-A2.
 XX 09-MAY-2001.
 XX 27-OCT-2000; 2000EP-00309475.
 XX 29-OCT-1999; 99JP-00309616.
 XX (RIKE) RIKEN KK.
 XX Yokoyama S, Hoseki J, Yano T, Koyama Y, Kuramitsu S;
 PI Kagamiyama H;
 XX WPI; 2001-337112/36.
 XX Novel mutant kanamycin nucleotidyltransferase useful for screening
 PT thermophilic bacteria such as Thermus thermophilus, has point mutations
 PT at specific positions and improved thermostability.
 XX Claim 2; Page 14-15; 30pp; English.
 XX The invention relates to mutant kanamycin nucleotidyltransferases (KNT)
 CC having improved thermostability. The mutant has one or more point
 CC mutations selected from Met57Leu, Ala62Val, Ser94Pro, Ser203Pro,
 CC Asp206Val, His207Gln, Ser220Pro, Ile234Val and Thr238Ala as against a
 CC WT-KNT (AAB62512). A KNT gene encoding the mutant can be used as a
 CC selective marker for thermophilic bacteria. The marker can be used for
 CC screening thermophilic bacteria, preferably Thermus thermophilus. The
 CC present sequence represents a S. aureus mutant KNT protein KT3-11
 XX
 SQ Sequence 253 AA;

Alignment Scores:
 Pred. No.: 1,12e-124 Length: 253
 Score: 1267.00 Matches: 240
 Percent Similarity: 96.4% Conservative: 4
 Best Local Similarity: 94.9% Mismatches: 9
 Query Match: 23.0% Indels: 0
 DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x AAB62513 (1-253)

QY 1579 GTCAATGACCAATAATGACTAGAGAAAGATGAAGATTTGTCATGAATTAAG 1638
 DB 1 MetLysGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
 QY 1639 GAACGAATATGGATAATATGGGATGATGTTAAGCTATTGGTGTATGGCTCTCTT 1698
 DB 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTGCTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
 DB 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
 QY 1759 GAAGCAGAGTTTACGCCATGAATGGACACCGGTGAGTGGAGGTGGAAGTGAATTTGAT 1818
 DB 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
 QY 1819 AGCGAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878
 DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100
 QY 1879 GGTCAATTTTCTCTATTTCGCGATTATGATTCAGGTGGATPACTTTAGAGAAAGTGTAT 1938
 DB 101 GlyArgPheSerIleLeuProIleTyrAspProGlyTyrPheGluLysValTyr 120
 QY 1939 CAAACTCTAAATCGGTAGAGCCCAACGTTCCACGATCGGATTTGTGCGCTTATCGTA 1998
 DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
 QY 1999 GAAGAGCTGTTGAATATGACGAGCAAAATGGCGTAATATTTCGTGCAAGGACGACAAACA 2058
 DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
 QY 2059 TTTCTACCATCTTGTACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCAT 2118
 DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
 QY 2119 CATCGCATCTGTTATACGACGCGCTTCGTTTAACTGAAGCAGTGAAGCAATCAGAT 2178
 DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
 QY 2179 CTTCTCTCAGGTTATGACCATCTGTCAGTTCGTAATGTCGTCAACTTTCGACTCT 2238
 DB 201 LeuProSerGlyTyrAspHisLeuCysGlnLeuValMetSerGlyGlnLeuSerAspSer 220
 QY 2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTCGACAGAACGA 2298
 DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
 QY 2299 CACGGATATAGTGGATGTCGAAACGATACCATTT 2337
 DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 8
 AAB62514
 ID AAB62514 standard; protein; 253 AA.
 XX AAB62514;
 XX 23-JUL-2001 (first entry)
 XX Mutant kanamycin nucleotidyltransferase HTK.
 XX Kanamycin nucleotidyltransferase; KNT; thermostability; mutation; mutant;
 KW marker; thermophilic; Thermus thermophilus.
 XX Staphylococcus aureus.
 OS Synthetic.
 XX EP1097990-A2.
 XX 09-MAY-2001.
 XX 27-OCT-2000; 2000EP-00309475.
 XX 29-OCT-1999; 99JP-00309616.
 XX (RIKE) RIKEN KK.
 XX Yokoyama S, Hoseki J, Yano T, Koyama Y, Kuramitsu S;
 PI Kagamiyama H;


```
QY 2373 GTGAATAAGGAACGCTGCTTGAAGCGCGCAAGCGCGCACGATCTGAATATCGGCACA 2432
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 86 IleThrIysGluGluIleLeuAlaGlyAlaLysArgAlaPheGluAenLysIleGlyThr 105

QY 2433 TATTGTATCTGGCAGCGCGCAGAGTCCCTCTAAACAGAGAAGTGAATCAGGTCTGATAT 2492
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 106 TyrCysIleValAlaSerGlyArgGlyProThrArgLysAspValAenValSerGlu 125

QY 2493 GCGGTTTCAGGAATTAAGAGACCTGATGAGCTGAAGATTGTGCATGCTCTTGCACTGTTG 2552
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 126 AlaValGluGluIleLysAlaLysTyrGlyLeuLysValCysAlaCysLeuGlyLeuLeu 145

QY 2553 AAGCCAGCAGCGCGCAGCGCTCAAAAGATGCAGGAGTAGACCGCTATATCAATATTTG 2612
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 146 LysGluGluGlnAlaGlnLeuLysGluAlaGlyValAspArgTyrAenHisAenLeu 165

QY 2613 AATACGTACAGAGAAACCACTCAACATCAACACCTCATACATACAGATGACAGATC 2672
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 166 AsnThrSerGluArgHisHisSerTyrIleThrThrHisThrTyrGluAspArgVal 185

QY 2673 AATACGTTCAATCGCAAGAAATCGGGCTGCTCCGTTGTCAGCGCCATTATCGGG 2732
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 186 AsnThrValGluValValLysHisGlyIleSerProCysSerGlyAlaIleIleGly 205

QY 2733 ATGAAGGAGACGAAACAGGATGTCATTGACATCGCCAAAGCTTGAAGGCTCTTGACGCG 2792
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 206 MetLysGluThrLysMetAspValValGluIleAlaArgAlaLeuHisGlnLeuAspAla 225

QY 2793 GATTCCATTCCTGTAATTTTTCATGCAATTTGATGGCAGCGCGTTAGAACGGCTCAAC 2852
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 226 AspSerIleProValAenPheLeuHisAlaIleAspGlyThrLysLeuGluGlyThrGln 245

QY 2853 GAATTAACCCGCTGATTTGTTAAAGTCTCGCGCTGTTCCGTTTATCAATCCATCA 2912
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 246 AspLeuAenProArgTyrCysLeuLysValLeuAlaLeuPheArgTyrMetAenProser 265

QY 2913 AAAGAAATTCGCATTTCCGAGGAAAGAGAGTCAATCTCCGCACATTCAGCCATTAGGG 2972
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 266 LysGluIleArgIleSerGlyArgGluValAenLeuGlyPheLeuGlnProPheGly 285

QY 2973 CTTTACGCCCAACTCCATTTTGTTCGAGAGACTACTTAAACAACCTGCCCGGCGAAGAGAG 3032
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 286 LeuTyrAlaAlaAenSerIlePheValGlyAspTyrLeuThrThrGluGlyGlnGluAla 305

QY 3033 ACGGAGGATCAATAAATGCTGAGTATTTAGGCTTTGAAGTTGAATCAGTCAAGAAATG 3092
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 306 AsnSerAspTyrArgMetLeuGluAspLeuGlyPheGluIleGluLeuThrGlnLysGln 325

QY 3093 AAG 3095
   ::
Db 326 Glu 326

RESULT 10
ID AAP81191
AC standard; protein; 337 AA.
XX
AC AAP81191;
XX
XX
DT 21-OCT-2004 (revised)
DT 25-MAR-2003 (revised)
DT 04-DEC-1990 (first entry)
XX
DE Biotin synthetase (Bio B) of B.sphaericus IFO 3525.
XX
KW biotin biosynthesis; fermentative biotin production; plasmid pTG1400;
KW biotin synthetase; Bio B; seborrhoeic dermatitis treatment.
XX
OS Bacillus sphaericus.
OS Unidentified.
XX
PN EP26240-A.
XX
```

```
PD 04-MAY-1988.
XX
XX 28-SEP-1987; 87BP-00402157.
XX
PR 30-SEP-1986; 86FR-00013603.
PR 18-MAY-1987; 87FR-00006916.
XX
PA (TRGE ) TRANSGENE SA.
XX
PI Gloeckler, Speck D, Lemoine Y;
XX
XX WPI; 1988-121306/18.
DR N-PSDB; AAN81546.
XX
XX New DNA sequences encoding enzymes involved in biotin biosynthesis -
PT isolated from Bacillus sphaericus, and transformants useful in
PT fermentative biotin prodn.
XX
PS Disclosure; Page: 51pp; French.
XX
CC DNA was isolated from B.sphaericus IFO 3525, cut with HindIII and the
CC fragments sub-cloned into pBR322. The recombinant plasmids were used to
CC transform E.coli bio- mutants. Plasmid pTG1400 was present in one clone
CC (E.coli C600; CNCM I-608) which showed complementation for bio A and bio
CC D. The plasmid contains a 4.3kb insert which contains 4 long open reading
CC frames (IORFs) including that which encodes Bio B. This enzyme is well
CC characterised. A second clone was isolated which contained DNA encoding
CC bio F and bio C. The insert present in the plasmid of this clone and the
CC 4.3kb insert from pTG1400 were cloned together in pBR322 to produce a
CC plasmid (pTG1440) which complements biotin auxotrophy in a bio delta FCD
CC mutant. Transformed cells are used to produce biotin by culture in a
CC medium contg pimelic acid and/or biotin vitamer. Co-culture of two
CC transformants is preferred where one synthesises the vitamer from pimelic
CC acid and the other converts the vitamer to biotin. See also AAN81551.
CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
CC correct PI field.)
CC
CC Revised record issued on 21-OCT-2004 : Correction to location
XX
XX Sequence 337 AA;
SQ
Alignment Scores:
Pred. No.: 1,19e-91 Length: 337
Score: 957.00 Matches: 180
Percent Similarity: 84.9% Conservative: 33
Best Local Similarity: 71.7% Mismatches: 38
Query Match: 17.3% Indels: 0
DB: 1 Gaps: 0
US-10-681-086-1 (1-3156) x AAP81191 (1-337)
QY 2343 TCGAAGCGCGCATGCTGAGTCTTACCGGATCGTGAATAAGGAAACGCTTGAAGCGCG 2402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 SerThrAlaProIleGluLysTyrProPheIleThrLysGluGluIleLeuAlaGlyAla 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2403 AAGCGCGCGCACGATCTGAATATCGGCACATATTGTATCGTGCAGCGCGCAGAGTCCG 2462
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 LysArgAlaPheGluAenLysIleGlyThrTyrCysIleValAlaSerGlyArgGlyPro 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2463 TCTAACAGAGAAGTGAATCAGGTCTGATAGTCGGTTTCAGGAATTTAAAGAGAGTATGGA 2522
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 122 ThrArgLysAspValAenValSerGluAlaValGluIleLysAlaLysTyrGly 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2523 CTGAAGATTGTGCATGCTTGGACTGTTGAACCCAGAGCGCGGAGCGCTCAAGAT 2582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 LeuLysValCysAlaCysLeuGlyLeuLeuLysGluGlnAlaGlnLeuLysGlu 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2583 GCAGGAGTAGACCGCTATAATCAATTAATTTCAATCGTCAAGAGAAACCATTTCAACATC 2642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 AlaGlyValAspArgTyrAenHisAenLeuAenThrSerGluArgHisHisSerTyrIle 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2643 ACAACCTCATACATACAGATGACAGATCAATACCGTTGAAATTCGAAAGAAATCGGG 2702
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 182 ThrThrThrGlnThrTyrGluAspArgValaAenThrValGluValValLysLysHisGly 201
Qy 2703 CTGTCCTCGCTCTCAGCGCCATTATCGGATGAGGAGACGAAACAGGATCTCATTCAC 2762
Db 202 IleSerProCysSerGlyAlaIleGlyMetLysGluThrLysMetAspValValGlu 221
Qy 2763 ATCCGCAAAAGCTTGAAGGCTCTGACCGGATTCCTGTAATTTTTCATGCA 2822
Db 222 IleAlaArgAlaLeuHisGlnLeuAspAlaAspSerIleProValaAenPheLeuHisAla 241
Qy 2823 ATTGATGCGCAGCGCTAGAGGCGTCAACGAAATTAACCGCGTGTATTGTTAAAGTG 2882
Db 242 IleAspGlyThrLysLeuGluGlyThrGlnAspLeuAenProArgTyrCysLeuLysVal 261
Qy 2883 CTGCGCGCTGTCCTTTTATCAATCCATCAAAAGAAATTCGCAATTCGCGAGGAGAG 2942
Db 262 LeuAlaLeuPheArgTyrMetAsnProSerLysGluIleArgIleSerGlyArgGlu 281
Qy 2943 GTCAATCTCCGACATTCAGCCATTAGGCGCTTTACGCGCAAACTCCATTTTGTGCGA 3002
Db 282 ValaAenLeuGlyPheLeuGlnProPheGlyLeuTyrAlaAlaAenSerIlePheValGly 301
Qy 3003 GACTACTTAAACACTGCGGCGAAGAGAGAGAGAGATCAATAAATGCTGAGTGATT 3062
Db 302 AspTyrLeuThrThrGluGlyGlnGluAlaAenSerAspTyrArgMetLeuGluAspLeu 321
Qy 3063 GGCTTGAAGTTGAATCAGTCGAGAAATGAAG 3095
Db 322 GlyPheGluIleGluLeuThrGlnLysGlnGlu 332

RESULT 11

AAW30520
ID AAW30520 standard; protein; 387 AA.
XX
AC AAW30520;
XX
XX 17-OCT-2003 (revised)
DT 26-OCT-1998 (first entry)
XX
XX Kurthia sp. KAPA synthetase.
DE
XX Biotin; vitamin; bioF gene; KAPA synthetase.
XX
XX Kurthia sp. 538-KA26; (DSM 10609).
XX
XX EP853127-A2.
XX
XX 15-JUL-1998.
XX
XX 18-SEP-1997; 97EP-00116237.
XX
XX 27-SEP-1996; 96EP-00115540.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
XX Furuichi Y, Hoshino T, Kimura H, Kiyasu T, Nagahashi Y;
XX
XX WPI; 1998-364652/32.
XX
XX N-PSDB; AAW42056.
XX
XX New DNA and vectors encoding polypeptides - used for recombinant
PT production of biotin.
XX
XX Claim 1; Page 20-21; 45pp; English.

XX This is the KAPA synthetase encoded by a newly isolated bioF gene (see
CC AAW42056) of Kurthia sp. 538-KA26 (DSM 10609). This enzyme catalyses the
CC conversion of pimelyl CoA to KAPA in the biosynthesis of biotin.
CC Polypeptides (see AAW30518-25) encoded by novel bioD, bioA, bioP, bioB,
CC bioH, bioFII, bioHII and bioC genes (see AAW42054-61) of Kurthia sp. are
CC provided, as well as vectors comprising one or more of the genes,
CC transformed cells, and a process for the production of biotin that
CC comprises cultivating the transformed cells and isolating biotin from the

CC culture medium. Biotin is used in the preparation of pharmaceutical, food
CC or feed compositions. The synthesis is an improvement on prior art
CC methods involving fermentation of e.g. E.coli and Bacillus sphaericus,
CC which have low productivity due to accumulation of DHB, a biotin
CC precursor. (Updated on 17-OCT-2003 to standardise OS field)

XX
SQ Sequence 387 AA;

Alignment Scores: 5.12e-90 Length: 387
Pred. No.: 942.00 Matches: 184
Score: 942.00
Percent Similarity: 73.6% Conservative: 81
Best Local Similarity: 51.1% Mismatches: 87
Query Match: 17.1% Indels: 8
Gaps: 2

US-10-681-086-1 (1-3156) x AAW30520 (1-387)

Qy 37 TGGTTAAACGAGCGGTTAGACAGAAATGAAAGAGCCGCGGTACATCGTAACCTCGCGTCA 96
Db 3 TtpGlulysGlu---LeuGluLysIleLysGluGlyGlyLeuTyrArgGlnLeuGlnThr 21
Qy 97 ATGATCGAGCGCGGTTCCAGAGAGGAAT-----ATTGATGGCGAAATCAACG 147
Db 22 ValGlu-----ThrMetSerAspGlnGlyTyrAlaMetValaAenGlyLysLysMetMet 39
Qy 148 GTCTGGTCTCAACAATTTAGGCGTCGACGAGCATAGACGTTTGCATCGATCGAGCC 207
Db 40 MetPheAlaSerAenAenTyrLeuGlyIleAlaAenAspGlnArgLeuIleGluAlaSer 59
Qy 208 CAAACAGCATTCGACCAATTTGGACAGAGCGGTTCCACGTTTAACTGACAGAGCAAT 267
Db 60 ValGlnAlaThrGlnArgPheGlyThrGlySerThrGlySerArgLeuThrThrGlyAsn 79
Qy 268 TCGGTCTGGCATGAAAGCTAGAAAAGAGATTGCCAGCTTTAACTGACAGAGCGGCC 327
Db 80 ThrIleValHisGluLysLeuGluLysArgLeuAlaGluPheLysGlnThrAspAlaAla 99
Qy 328 CTGCTGTTTCGAGCGGTTACTTGGCCAATGTCGGTCTCCTTTCATCTCCGCCAGAAAG 387
Db 100 IleValLeuAenThrGlyTyrMetAlaAenIleAlaAlaLeuThrThrLeuValGlySer 119
Qy 388 GAAGATGTCATTTTAAGTGACCCAGCTCAATCATGCAAGTATGATCGAGCGGCTCGGACTT 447
Db 120 AspAspLeuIleLeuSerAspGluMetAsnHisAlaSerIleIleAspGlyCysArgLeu 139
Qy 448 TCTAAGGCTGATACAGTTGTTATCGGCATATTGATATGATATGATCTTTGAAAACAAGCTG 507
Db 140 SerArgAlaGluThrIleIleTyrArgHisAlaAspLeuLeuAspLeuGluMetLysLeu 159
Qy 508 AATCAAAACACAGCGTTATCAGCGCGTTTATCGTAACAGCGGAGTATTCAGCATGGAT 567
Db 160 GlnIleAenThrArgTyrArgLysArgIleIleValThrAspGlyValPheSerMetAsp 179
Qy 568 GGCACAATCGCCCTCTTGATCAGATCATCTCACTTCGAAACCGCTATCATGCTCTCGTG 627
Db 180 GlyAspIleAlaProLeuProGlyIleValGluLeuAlaLysArgTyrAspAlaLeuVal 199
Qy 628 GTCGTTGATGATGCCACGCAAGAGGATTTGGCGGATTCGGGACAGAGAACAGAGTAA 687
Db 200 MetValAspAspAlaHisAlaThrGlyValLeuGlyLysAspGlyArgGlyThrSerGlu 219
Qy 688 TACTTTCGTTGTTGTCCTCC-----GACATTCTTATCGGCACCTTAAGCAAGCTGTTGCG 741
Db 220 HisPheGlyLeuLysGlyLysIleAspIleGluMetGlyThrLeuSerLysAlaValGly 239
Qy 742 GCGGAAGAGGATTTTGGCGGAGGATCAGCGGTCTTCATCGACTTTTGTCTGAACCATGCC 801
Db 240 AlaGluGlyGlyTyrIleAlaGlySerArgSerLeuValAspTyrValLeuAenArgAla 259
Qy 802 AGAATATTATCTTTAAACCGCTATTTCGCGCAGCCAGCTGTCGGGCTGCTCAGAGGCT 861
Db 260 ArgPropheValPheSerThrAlaLeuSerAlaGlyValValAlaAlaSerAlaLeuThrAla 279

[illegible]

Best Local Similarity:	58.2%	Mismatches:	51
Query Match:	15.4%	Indels:	18
DB:	2	Gaps:	1
 US-10-681-086-1 (1-3156) x AAW30521 (1-338)			
QY	2307	TATAGTGGATGTCTCAAAACGCATACATTTTGAAATCGAAA-----	2348
DG	60	TyrGlyLysLysValLysLeuAenMetIleAenThrLysSerGlyLeuCysProGlu	79
QY	2349	-----GCCGCCATTGAGTCTTACC GGATG	2372
DB	80	AspCysGlyTyrCysSerGlnSerIleValSerGluAlaProIleAspLysfTyAlaTrp	99
QY	2373	GTGAATAAGGAAACGGCTGTTGAAGCGCGAACGGCGCACCATCTGTAATATCGGCACA	2432
DB	100	LeuThrLysGluLysIleValGluGlyAlaGlnGluSerIleArgArgLysAlaGlyThr	119
QY	2433	TATTGTATCTGTCGAACGGCGCAGAGTCCGTCTAACAGAGAAGTGGATCAGGTCGTAGAT	2492
DB	120	TyrCysIleValAlaSerGlyAsGProThrAsnArgGluIleAspHisValIleGlu	139
QY	2493	CGCGTTCAGAAATTAAGAGAGCATGACTCAAGATTGTGCATGCTCTTGACATGTTG	2552
DB	140	AlaValLysGluIleArgGluThrThrAspLeuLysIleCysCysLeuGlyPheLeu	159
QY	2553	AAGCCAGAGCAGCGAAGCGCTCAAAGATGACGAGTAGACCGCTATATCATATAATTG	2612
DB	160	AsnGluThrHisAlaSerLysLeuAlaGluAlaGlyValHisArgTyrLysHisAsnLeu	179
QY	2613	AATACGTCACAGAGAACCATTCAAAATCACAACCTCACATACATACGATGACAGATC	2672
DB	180	AsnThrSerGlnAspAsnTyrLysAsnIleThrSerThrHisThrTyrGluAspA-gVal	199
QY	2673	AATACGTTGAAATCGCAAAAGAAATCGGGCTCTCCGGTGTTCAGCGCCATATCGGG	2732
DB	200	AspThrValGluAlaValLysGluAlaGlyMetSerProCysSerGlyAlaIlePheGly	219
QY	2733	ATGAAGGAGACGAAACAGGATGTCATTGACATCGCCAAAGCTTTGAAGGCTCTTGACGG	2792
DB	220	MetAsnGluSerAsnGluGluAlaValGluIleAlaLeuSerLeuArgSerLeuAspAla	239
QY	2793	GATTCCATTCTCTGTAATTTTTTGTCATGCAATTGATGGCAGCGCGTTAGAAGCGCTCAAC	2852
DB	240	AspSerileProCysAsnPheLeuAenAlaIleAspGlyThrProLeuGluGlyThrSer	259
QY	2853	GAATTAAACCGCTGATTTGTTTAAAGTCTCGGCGTGTTCGGTTTATCAATCATCA	2912
DB	260	GluLeuThrProThrLysCysLeuLysLeuIleSerMetMetArgPheValAsnProSer	279
QY	2913	AAAGMAATTCGATTTCCGGAGACGAGGTCAATCTCCGACATTTGCAGCCATTAGGG	2972
DB	280	LysGluIleArgLeuAlaGlyGlyArgGluValAsnLeuArgSerMetGlnProMetAla	299
QY	2973	CTTTACGCCCAACTCCATTTTGTTCGGAGACTACTTAACAACTCGCGCGCAAGAGAG	3032
DB	300	LeuTyrAlaAlaAsnSerIlePheValGlyAspTyrLeuThrThrAlaGlyGlnGluPro	319
QY	3033	ACGAGAGGATCAATAAATGCTGAGTGATTTAGCGTTTGAAGTTCAA	3077
DB	320	ThrAlaAspTrpGlyIleIleGluAspLeuGlyPheGluIleGlu	334
 RESULT 14			
ID	ADFA1704	standard; protein; 392 AA.	
XX	ADFA1704;		
XX	AC		
XX	DT	12-FEB-2004 (first entry)	
XX	DE	Bacillus subtilis Kbl amino acid sequence SEQ ID NO:26.	
XX	KW	Bacillus; sbo; slr; ybcO; csn; spollSA; sigB; phrC; rapA; CseS; trpA;	

RESULT 15
ADS44705
ID ADS44705 standard; protein; 392 AA.
XX
AC ADS44705;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #23135.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 23135, 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 392 AA;

Alignment Scores:

Pred. No.: 1.71e-79 Length: 392

```

Db      327 IleGlyAspGluGlyValAlaLysGlnPheSerAspGlnLeuLeuSerArgGlyValPhe 346
      |||||||
      |||:~::~:~::|||
Qy      1051 GCTCTGCAATTCGGCCGCCAACCCTTCGCCCGGGTGAAGCCGGATTGGAAGCTTGGGC 1110
      |||:~::~:~::|||
Db      347 AlaGlnSerIleValPheProThrValAlaLysGlyLysAlaArgIleArgThrIleIle 366
      |||:~::~:~::|||
Qy      1111 AGCAGGTCGAGATCAGGGAATGAG 1134
      |||:~::~:~::|||
Db      367 ThrAlaGluHisThrLysAspGlu 374
      |||:~::~:~::|||

```

Search completed: February 10, 2006, 02:33:17
 Job time : 485.5 secs

This Page Blank (uspio)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2006, 02:33:31 ; Search time 21.4 Seconds

(without alignments)
2837.945 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggttacgacg.....atcggtatgatgtcgaaattc 3156

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abses/ABSSWEB_spool/US10681086/runat_09022006_140517_11758/app_query.fasta_1
-DB=PIR -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=550 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10681086 @CGN_1_1_63 @runat_09022006_140517_11758 -NCPU=6 -ICPU=3
-NO MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOF=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1836	33.3	389	2 F69594	8-amino-7-oxononan
2	1350	24.5	256	2 S09565	kanamycin nucleoti
3	1327	24.1	253	2 B24456	kanamycin nucleoti
4	1317	23.9	335	2 D69594	biotin synthase (E
5	1103	20.0	333	2 D83868	biotin synthase bi
6	971	17.6	332	1 J50274	biotin synthase (E
7	929	16.8	389	1 JQ0512	8-amino-7-oxononan
8	843	15.3	392	2 D69647	glycine C-acetyltr
9	793	14.4	335	2 B90044	biotin synthase [i
10	712	12.9	395	2 C84138	8-amino-7-oxononan
11	698.5	12.7	372	2 A64642	8-amino-7-oxononan
12	692	12.5	398	2 F75017	probable glycine C
13	688	12.5	386	2 AE1853	8-amino-7-oxononan
14	687	12.5	383	2 H89822	hypothetical prote

15	686	12.4	398	2 E71454	probable glycine C
16	676.5	12.3	373	2 G70355	8-amino-7-oxononan
17	667	12.1	437	2 S75730	8-amino-7-oxononan
18	651.5	11.8	424	2 A75286	probable glycine C
19	610	11.1	403	2 AD0008	glycine C-acetyltr
20	609	11.0	397	2 C82405	glycine C-acetyltr
21	601	10.9	398	2 AI0973	glycine C-acetyltr
22	598	10.8	398	1 XUECGA	glycine C-acetyltr
23	598	10.8	398	2 H86037	glycine C-acetyltr
24	598	10.8	398	2 G91190	2-amino-3-ketobuty
25	595	10.8	640	2 T34316	transferase - Stre
26	589.5	10.7	401	2 G83582	8-amino-7-oxononan
27	582	10.5	440	2 G87444	8-amino-7-oxononan
28	578.5	10.5	384	2 A82239	8-amino-7-oxononan
29	567.5	10.3	771	2 F70701	probable bioF2 pro
30	565.5	10.3	385	2 AI0596	8-amino-7-oxononan
31	560.5	10.2	380	2 D64129	probable 8-amino-7
32	559	10.1	378	2 AG3606	8-amino-7-oxononan
33	552.5	10.0	380	2 H81830	8-amino-7-oxononan
34	544.5	9.9	423	2 F97672	5-aminolevulinic a
35	544	9.9	414	2 A71646	5-aminolevulinat
36	543	9.8	405	2 S15996	5-aminolevulinat
37	542.5	9.8	380	2 C81194	8-amino-7-oxononan
38	540.5	9.8	332	2 E70384	biotin synthase (E
39	539	9.8	383	2 AF0141	8-amino-7-oxononan
40	538	9.8	386	2 C70540	probable 8-amino-7
41	536	9.7	384	2 F90735	8-amino-7-oxononan
42	536	9.7	558	2 T38542	probable 5-aminole
43	535	9.7	384	1 SYECKP	8-amino-7-oxononan
44	535	9.7	384	2 G85585	8-amino-7-oxononan
45	534.5	9.7	409	1 SYZJAL	5-aminolevulinat

ALIGNMENTS

RESULT 1

F69594

8-amino-7-oxononanoate synthase (BC 2.3.1.47) bioF - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C;Accession: F69594

R;Kunze, F.; Ogaawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Seginchi, J.; Sekowaka, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yamumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69594

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-389 <KUN>

A;Cross-references: UNIPROT:P53556, UNIPARC:UPI0000060924; GB:Z99119; GB:AL009126; NID:G

A;Experimental source: strain 168

C;Genetics:

A;Gene: bioF

C;Superfamily: 5-aminolevulinat synthase; glycine C-acetyltransferase homology
C;Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate
F;44-377/Domain: glycine C-acetyltransferase homology <GCA>
F;237/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:

Pred. No.: 9.13e-133

Score: 1836.00

Length:

Matches:

Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 33.3% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x F69594 (1-389)

```
Qy 22 TTGAAGATTGATCTCTGTTAAACAGCGGTTAGACAGAAAGAGCGCGGTACAT 81
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 1 MetLysIleAaspSerIrrPleuanGluargLeuAaspArgMetLysGluAalGlyValHis 20
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 82 CGTAACCTCGCGTCAATGATGAGCGCGCGGTTCCAGAGAGAAATATTGATGGCGAAAT 141
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 21 ArghenLeuAargSerMetAaspGlyAlaProValProGluAargAenIleAaspGlyGluAen 40
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 142 CAACGGTCTGCTCTCAAAACAAATATTATTAGGCTCGCAAGCATAGACGTTTGATCAT 201
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 41 GlnThrValTrrpSerSerAasnAenTyrLeuGlyLeuAlaSerAaspArgLeuIleAasp 60
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 202 GCAGCCCAACAGCATTCAGCAATTTGGGACAGGAAGCAGCGGTTACAGTTTAACGACA 261
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 262 GGCATTCGCGTCTGCGATGAAAGCTAGAAAAGAGATTGCCAGCTTTTAAACTGCAGAA 321
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 81 GlyAenSerValTrrpHisGluLysLeuGluLysLysIleAlaSerPheLysLeuThrGlu 100
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 322 GCGGCGCTGCTGTTTCAGCGGTTACTTGGCCAAATGTCGGTGTCCTTTTCCTTCGCCA 381
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAenValGlyValLeuSerSerLeuPro 120
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 382 GAAAGGAGAGATGTCATTTTAAGTCACAGCTCAATCATGCAAGTATCATCGAGCGGTCG 441
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 121 GluLysGluAaspValIleLeuSerAaspGlnLeuAenHisAlaSerMetLysAaspGlyCys 140
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 442 CGACTTTCTAAGGCTGATACAGTGTGTTTATCGGCATATTGATGAATGATCTTGAAAC 501
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 141 ArgLeuSerLysAlaAaspThrValValTyrArgHisIleAaspMetAasnAaspLeuAen 160
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 502 AAGTGAATGAAACACACGGTTATCAGCGCGTATTATCGTAACAGACGGAGTATTCAGC 561
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 161 LysLeuAenGluThrGlnArgTyrGlnArgArgPheIleValThrAaspGlyValPheSer 180
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 562 ATGATGCAACAATCGCCCTCTTGATCAGATCATCTCACTTGGCAACGCTATCATGCC 621
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 181 MetAepGlyThrIleAlaProLeuAaspGlnIleSerLeuAlaLysArgTyrHisAla 200
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 622 TTCGTGTCGTGTGATGATGCCCAACAGAGTTTTGGCGGATTCGGGCAACAGGAACG 681
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 201 PheValValValAaspAaspAlaHisAlaThrGlyValLeuLeuGlyAaspSerGlyGlnGlyThr 220
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 682 AGTGAATACTTGTGTTGTCGCCGACATTTGTTATCGGCACCTTAAGCAAAAGCTGTTGGC 741
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 221 SerGluTyrPheGlyValCysProAaspIleValIleGlyThrLeuSerLysAlaValGly 240
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 742 GCGGAAGAGGTTTTGCGCGAGGATCAGCGGCTCTTCATCGACTTTTCTGCAACCATGCC 801
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAaspPheLeuLeuAenHisAla 260
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 802 AGAACATTTATCTTTCAAAACCGCTATTCCGCCAGCCAGCTGTGCGGCTGCTCAGCAGGCT 861
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaAlaHisGluAla 280
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 862 TTCAACATCATTAAGCCAGCGGGGAAAAACAGACAGCTTTTATTCTTATATCATCGCATG 921
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 281 PheAenIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 922 ATCAGAACCGAGTCTCAAGAAATATGGTTATGTGTAAGAGGAGATCACACACCATATTATT 981
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 301 IleArgThrSerLeuLysAasnMetGlyTyrValValLysGlyAepHisThrProIleIle 320
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 982 CCTGTAGTCAITGGCGATGCCCATAAAAACGGTCTCTATTGCTTGAAAAAACTCGACGGCAAG 1041
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
Db 321 ProValValIleGlyAaspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
Qy 1042 GGAATTTTATGCTCTGTCATTCGCGCCCAACCTTTCGCGCGGTGAAGCGCGATTGCA 1101
Db 341 GlyIleTyrAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 360
```

RESULT 2

S09565
kanamycin nucleotidyltransferase (EC 2.7.7.-) [imported] - Staphylococcus aureus plasmid
A:Alcarnate names: emycin resistance protein; kanamycin inactivating enzyme; kanamycin
C:Species: Staphylococcus aureus
C:Date: 19-Mar-1997 #sequence revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S09565; T44127; A24456
R:Minton, N.P.; Swinfield, T.J.; Brehm, J.K.; Oultram, J.D.
Nucleic Acids Res. 18, 1651, 1990
A:Title: The gram-positive cloning vector pBD64 arose by a 1844 bp deletion of pC194 der
A:Reference number: S09565; MUID:90221915; PMID:2326208
A:Accession: S09565
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-256 <MIN>
A:Cross-references: UNIPROT:O87369; UNIPARC:UPI00000D474F; EMBL:X51450; NID:957998; PID:
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1990
R:Ito, T.; Katayama, Y.; Hiramatsu, K.
Antimicrob. Agents Chemother. 43, 1449-1458, 1999
A:Title: Cloning and nucleotide sequence determination of the entire mec DNA of pre-meth
A:Reference number: 222733; MUID:99278010; PMID:10348769
A:Accession: T44127
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-256 <ITO>
A:Cross-references: UNIPARC:UPI00000D474F; EMBL:D86934; PIDN:BAA82229.1
A:Experimental source: strain N315
R:Matsumura, M.; Katakura, Y.; Imanaka, T.; Aiba, S.
J. Bacteriol. 160, 413-420, 1984
A:Title: Enzymatic and nucleotide sequence studies of a kanamycin-inactivating enzyme en
A:Reference number: A24456; MUID:85006820; PMID:6090428
A:Accession: A24456
A:Molecule type: DNA
A:Residues: 1-5-256 <MAT>
A:Cross-references: UNIPARC:UPI000000131C
C:Genetics:
A:Gene: aadD
A:Genome: plasmid pBD64
C:Superfamily: Staphylococcus kanamycin nucleotidyltransferase
C:Keywords: antibiotic resistance; nucleotidyltransferase

Alignment Scores:
Pred. No.: 1.68e-95 Length: 256
Score: 1350.00 Matches: 256
Percent Similarity: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 24.5% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x S09565 (1-256)

```
Qy 1570 ATGAGATAGTGAATGACCAATTAATGCTAGACAGAAAGAAATGATGTTTCAT 1629
Db 1 MetArgIleValAenGlyProIleIleMetThrArgGluGluArgMetLysIleValHis 20
Qy 1630 GAAATTAAGCAACCAATATTGGATAAATATGCGGATGATGTTAAGGCTATTGGTGTATTAT 1689
Db 21 GluIleLysGluAargIleLeuAaspLysTyrGlyAaspValLysAlaIleGlyValTyr 40
Qy 1690 GGCTCTCTGTCGTCTCAGACTGATGGCCCTATTTCGATATATGAGATGATGTGTCTATG 1749
Db 41 GlySerLeuGlyArgGlnThrAaspGlyProTyrSerAaspIleGluMetCysValMet 60
Qy 1750 TCAACAGAGGACAGAGTTCCAGCCATGATGGCAACCGGTGAGTGGAAAGCTGGAAGCTG 1809
Db 61 SerThrGluGluAlaGluPheSerHisGluThrThrGlyGluThrLysValGluVal 80
```


QY 1810 AATTTCATGACGAGAGATTCTACTAGATTATGCAATCTCAGGTGGGAATCAGATTGGCCG 1869
Db 81 AsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100
QY 1870 CTTACACATGGTCAATTTTCTCTATTTCGCCGATTTTATGATTCAGGTGATCTAGAG 1929
Db 101 LeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120
QY 1930 AAAGTGTATCAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTGTGCC 1989
Db 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140
QY 1990 CTTATCTAGAGAGCTGTTTGAATATGACGAGCAAAATGGCGTAAATATCTGTGCAAGGA 2049
Db 141 LeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAenIleAglValGlnGly 160
QY 2050 CCGACACATTCTTACCATCTGTTACGAGTACAGGTAGCAATGCCAGTGCCTGATT 2109
Db 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180
QY 2110 GGTCTGCATCATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAG 2169
Db 181 GlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLys 200
QY 2170 CAATCAGATCTTCTTTCAGGTTATGACCATCTGTGCCAGTTCGTAATGTCTGGTCAACTT 2229
Db 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220
QY 2230 TCCGACTCTGAGAACTTCGGAATCGCTAGAGATTCGGAATTCGGAATTCAGGATGCG 2289
Db 221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAenGlyIleGlnGluTrp 240
QY 2290 ACAGAACGACACGGATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 ThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 256

RESULT 3
B24456
kanamycin nucleotidyltransferase (EC 2.7.7.-) - Bacillus sp. plasmid pTB913
N/Alternate names: kanamycin inactivating enzyme; kanamycin resistance protein
C/Species: Bacillus sp.
C/Date: 31-Dec-1990 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
R/Matsumura, M.; Katakura, Y.; Imanaka, T.; Aiba, S.
J. Bacteriol. 160, 413-420, 1994
A/Title: Enzymatic and nucleotide sequence studies of a kanamycin-inactivating enzyme
A/Reference number: A24456; MUID:85006820; PMID:6090428
A/Accession: B24456
A/Molecule type: DNA
A/Residues: 1-253 <MA1>
A/Cross-references: UNIPROT:P05058; UNIPARC:UPI000012DC09; GB:X03409; GB:K02552; NID:940
A/Molecule type: protein
A/Residues: 1-7 <MA2>
A/Cross-references: UNIPARC:UPI00001790DB
R/van der Lelie, D.; Bron, S.; Venema, G.; Oskam, L.
Nucleic Acids Res. 17, 7283-7294, 1989
A/Title: Similarity of minus origins of replication and flanking open reading frames of
A/Reference number: S05980; MUID:90016790; PMID:2677995
A/Accession: S05985
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 'NRIV', 2-253 <VAN>
A/Cross-references: UNIPARC:UPI000016E92D; EMBL:X15670; NID:940078; PIDN:CAA33715.1; PII
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1989
C/Genetics:
A/Gene: knl
A/Genome: plasmid
A/Start codon: GTG
C/Superfamily: Staphylococcus kanamycin nucleotidyltransferase
C/Keywords: antibiotic resistance; nucleotidyltransferase
F:1-253/Product: kanamycin nucleotidyltransferase #status experimental <MAT>

Alignment Scores:
Pred. No.: 9.78e-94 Length: 253
Score: 1327.00 Matches: 251
Percent Similarity: 99.6% Conservative: 1
Best Local Similarity: 99.2% Mismatches: 1
Query Match: 24.1% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x B24456 (1-253)

QY 1579 GTCAATGGACCCATTAATAGTACGAGAGAAAGATGAAGATTGTTTCATGAAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATGTGATAAAATATGGGATGATGTTAAAGCTATTGGTGTATTGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTCTGCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTGCATGTCACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGACAGTTTCAGCCATGAATCGACAAACCGGTGAGTGGAAAGTGGAAGTGAATTTTGAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluAlaAsnPheAsp 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTCGCCGATTTATGATTCAGGTGGAATCTTAGAGAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAACTCTTAATCGGTAGAGCCCAACGTTCCAGCATCGGATTTGTCGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTTCAATATGACGCAAAATGCGTAATATTCTGTGCAAGGACCCACACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTCTGATCTGACAGGTAGCAATGGCAGTGCCTATGTTGATTTGTTGCTGAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACCTGAACGAGTTAGCATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTTCCTTCAGGTTATGACCATCTGTGCGAGTTCTGTAATGTCTGTCAACTTTCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTTCTGGAATCGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 4
D69594
biotin synthase (EC 2.8.1.6) bioB [similarity] - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: D69594
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis. A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A:Reference number: A69590; MUID:98044033; PMID:9384377

A:Accession: D69594
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-335 <KUN>
A:Cross-references: UNIPROT:P53557; UNIPARC:UPI0000060922; GB:Z99119; GB:AL009126; NID:9
C:Experimental source: strain 168
C:Genetics:
C:Superfamily: biotin synthetase
C:Keywords: 2Fe-2S; biotin biosynthesis; iron-sulfur protein; metalloprotein; sulfuran
P:65,69,72,201/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Alignment Scores:
Pred. No.: 6.13e-93 Length: 335
Score: 1317.00 Matches: 263
Percent Similarity: 92.0% Conservative: 2
Best Local Similarity: 91.3% Mismatches: 5
Query Match: 23.9% Indels: 18
DB: 2 Gaps: 1

US-10-681-086-1 (1-3156) x D69594 (1-335)

QY 2307 TATAGTGGATGTGTCAAACGACATACCATTTTGAAT----- 2342
DB 48 TyrGlyLysValLeuAsnMetIleMetAenAlaLysSerGlyLeuCysProGlu 67
QY 2343 -----TCGAAAGCGCGATTTGAGTCTTACCGGATG 2372
DB 68 AsnCysGlyTyrCysSerGlnSerAlaIleSerLysAlaProIleGluSerTyrArgMet 87
QY 2373 GTGAATAAGGAACGCTCTTGAAGCGCGAGCGGCGACGATCTGAATATCGGCACA 2432
DB 88 ValAsnLysGluThrLeuLeuGluGlyAlaLysArgAlaHisAspLeuAsnIleGlyThr 107
QY 2433 TATGCTATCGTGGCAGCGGAGAGTCCGCTTAACAGAGAGTGGATCAGTCTGTAGAT 2492
DB 108 TyrCysIleValAlaSerGlyArgGlyProSerAenArgGluValAspGlnValAsp 127
QY 2493 GCGGTTTACGAGAAATTAAGAGACGATATGAGTCTTGAAGATTTGCTTGGACTGTG 2552
DB 128 AlaValGlnGluIleLysGluThrTyrGlyLeuLysIleCysAlaCysLeuGlyLeuLeu 147
QY 2553 AAGCAGACGAGCGGAGCGGCTCAAGATCGAGAGTAGACCGCTATAATCAATAATTG 2612
DB 148 LysProGluGlnAlaLysArgLeuLysAspAlaGlyValAspArgTyrAsnHisAsnLeu 167
QY 2613 AATAGCTCACAGAGAACCATTCAAACATCAACCTCACATACATACGATGACAGATC 2672
DB 168 AsnThrSerGlnArgAsnHisSerAsnIleThrThrSerHisThrTyrAspAspArgVal 187
QY 2673 AATACGCTTGAATCGCAAAAGATTCGGGCTGCTCGGTTTACGAGCGGCATTTACGG 2732
DB 188 AsnThrValGluIleAlaLysGluSerGlyLeuSerProCysSerGlyAlaIleGly 207
QY 2733 ATGAAGAGACGAAACAGAGATGTCATTCATCGCCAAAGCTTGAAGCGCTTTCAGCG 2792
DB 208 MetLysGluThrLysGlnAspValIleAspIleAlaLysSerLeuLysAlaLeuAspAla 227
QY 2793 GATCCATTCCTGTAATTTTTCGATCGAATTCGATGCGCAGCGCTTAGAGGCGCTCAAC 2852
DB 228 AspSerIleProValAsnPheLeuHisAlaIleAspGlyThrProLeuGluGlyValAsn 247

QY 2853 GAATTAACCCGCTGTATTGTTTAAAGTGTGGCGCTGTTCCGTTTTTATCAATCCATCA 2912
DB 248 GluLeuAsnProLeuTyrCysLeuLysValLeuAlaLeuPheArgPheIleAsnProSer 267
QY 2913 AAAGAAATTCGATTTCCGAGGAGAGAGTCAATCTCCGACATTCGACCCATTAGGG 2972
DB 268 LysGluIleArgIleSerGlyGlyArgGluValAsnLeuArgThrLeuGlnProLeuGly 287
QY 2973 CTTTACCGCCCAACTCCATTTTGTGGAGACTACTTAACAATCGCGGCGCAGAGGAG 3032
DB 288 LeuTyrAlaAlaAsnSerIlePheValGlyAspTyrLeuThrThrAlaGlyGlnGluGlu 307
QY 3033 ACGGAGGATCATAAATGCTCAGTGATTTAGGCTTTGAAGTTGAATCAGTCGAAGAAATG 3092
DB 308 ThrGluAspHisLysMetLeuSerAspLeuGlyPheGluValGluSerValGluGluMet 327
QY 3093 AAGCTAGTTTAAGTGCAGAAAGC 3116
DB 328 LysAlaSerLeuSerAlaLysSer 335

RESULT 5

biotin synthase bioB [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: D83868

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: D83868

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <STO>

A:Cross-references: UNIPROT:Q9KRC26; UNIPARC:UPI000000C3CA9; GB:AP001513; GB:BA000004; NID

A:Experimental source: strain C-125

C:Genetics:

A:Gene: bioB

C:Superfamily: biotin synthetase

Alignment Scores:
Pred. No.: 1.64e-76 Length: 333
Score: 1103.00 Matches: 206
Percent Similarity: 85.4% Conservative: 39
Best Local Similarity: 71.8% Mismatches: 24
Query Match: 20.0% Indels: 18
DB: 2 Gaps: 1

US-10-681-086-1 (1-3156) x D83868 (1-333)

QY 2307 TATAGTGGATGTGTCAAACGACATACCATTTTGAAT----- 2342
DB 47 TyrGlyLysValLeuAsnMetIleMetAenAlaLysSerGlyPheCysProGlu 66
QY 2343 -----TCGAAAGCGCGATTTGAGTCTTACCGGATG 2372
DB 67 AsnCysGlyTyrCysSerGlnSerIleSerLysAlaProIleAspAlaTyrProMet 86
QY 2373 GTGAATAAGGAACGCTCTTGAAGCGCGAGCGGCGACGATCTGAATATCGGCACA 2432
DB 87 ValAsnLysGluThrIleLeuGluGlyAlaLysArgAlaHisGluLeuAsnValGlyThr 106
QY 2433 TATTGATCTGGCAACGCGCAGAGTCCGCTTAACAGAGAGTGGATCAGTCTGTAGAT 2492
DB 107 TyrCysIleValAlaSerGlyArgGlyProThrAsnArgAspIleAspHisValThrGlu 126
QY 2493 GCGGTTTACGAGAAATTAAGAGACGATATGAGTCTTGAAGATTTGCTTGGACTGTG 2552
DB 127 AlaValArgGluIleLysAspThrTyrGlyLeuLysIleCysAlaCysLeuGlyIleLeu 146
QY 2553 AAGCAGACGAGCGGAGCGGCTCAAGATGCGAGAGTAGACCGCTATAATCAATAATTG 2612
DB 147 LysProGluGlnAlaGluGlnLeuLysAlaAlaGlyValAspArgTyrAsnHisAsnVal 166

A;Cross-references: UNIPROT:P22806; UNIPARC:UPI0000126987; GB:M29291; NID:q142592; PIDN:
A;Experimental source: strain IF03525
Biochem. J. 283, 327-331, 1992
A;Title: The 8-amino-7-oxopelarginate synthase from *Bacillus sphaericus*. Purification an
A;Reference number: S21284; MUID:92246854; PMID:1575677
A;Accession: S21284
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <P>O>
A;Cross-references: UNIPARC:UPI0000172399
C;Genetics:
A;Gene: bioF
C;Function:
A;Description: catalyzes the condensation of pimelyl-CoA and L-alanine to form 8-amino-7
C;Superfamily: 5-aminolevulinatase synthase; Glycine C-acetyltransferase homology
C;Keywords: acyltransferase; biotin biosynthesis; coenzyme A; phosphoprotein; pyridoxal
F;42-377/Domain: glycine C-acetyltransferase homology <GCA>
F;237/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:
Pred. No.: 3.88e-63 Length: 389
Score: 929.00 Matches: 187
Percent Similarity: 70.6% Conservative: 67
Best Local Similarity: 51.9% Mismatches: 100
Query Match: 16.8% Indels: 6
DB: 1 Gaps: 2

US-10-681-086-1 (1-3156) x JQ0512 (1-389)

```
Qy 40 TTAACGAGCGGTAGACAGA-----ATGAAGAAGCGCGGTACATCGTAAC 87
Db 1 MetAsnAspArgPheArgArgGluLeuGlnValleGluGluGlnGlyeThrArgLys 20

Qy 88 CTGCGGTCAATGGATGGAGCGCGGTCCAGAGAGGAATATTGATGGCGAAATCAACG 147
Db 21 LeuArgLeuPheSerThrGlyAsnGluSerGluValValMetAsnGlyLysLysPheLeu 40

Qy 148 GTCTGGTCTCAACAATATTATAGGCTCCAGCGATAGACGTTTGATCGATCGAGCC 207
Db 41 LeuPheSerSerAsnAsnTyLeuGlyLeuAlaThrAspSerArgLeuLysLysAla 60

Qy 208 CAACAGCATTTGGACAGGAAGAGCGGTTCACGTTTAACTGACAGAGCGGCC 267
Db 61 ThrGluGlyIleSerLysTyGlyThrGlyAlaGlySerArgLeuThrGlyAsn 80

Qy 268 TCGTCTGGCATGAAAGCTAGAAGAGATTGCCAGCTTTAACTGACAGAGCGGCC 327
Db 81 PheAspIleHisGluGlnLeuGluSerGluIleAlaAspPheLysLysThrGluAla 100

Qy 328 CTGCTGTTTCGAGCGGTACTTGGCCAAATGTCGGTGCCTTTCATCTTCCAGAAAG 387
Db 101 IleValPheSerSerGlyTyLeuAlaAsnValGlyValIleSerSerValMetLysAla 120

Qy 388 GAAGATGCTATTTAAGTGACAGCTCAATCATGCAAGTATGATCGAGCGTCCGACTT 447
Db 121 GlyAspThrIlePheSerAspAlaTrpAsnHisAlaSerIleIleAspGlyCysArgLeu 140

Qy 448 TCTAAGCGTGCATGATGTTTATCGGCATATTGATGATGAATGATCTTTGAAACAAGCTG 507
Db 141 SerLysAlaLysThrIleValTyGluHisAlaAspMetValAspLeuGluArgLysLeu 160

Qy 508 AATGAACACAGCGTTATCAGCGCGGTTTATCGTACAGAGCGAGTATTCAGGATGAT 567
Db 161 ArgGlnSerHisGlyAspGlyLeuLysPheIleValThrAspGlyValPheSerMetAsp 180

Qy 568 GGCACAAATCGCCCTCTTGATCAGATCTCACTTTCGAAACGCTATCATGCTTCGTG 627
Db 181 GlyAspIleAlaProLeuProlyIleValGluLeuAlaLysGlyTyLysAlaIle 200

Qy 628 GTCGTTGATGATGCCCAACGAGGATTTTGGCGGATTCGGGACAGGAGGAGTCAA 687
Db 201 MetIleAspAspAlaHisAlaThrGlyValLeuGlyAsnAspGlyCysGlyThrAlaAsp 220
```

```
Qy 688 TACTTTGGTGTGTTGTCCTCC-----GACATTTGTTATCGGCACCTTAAGCAAGCTGTGGC 741
Db 221 TyrPheGlyLeuLysAspGluIleAspPheThrValGlyThrLeuSerLysAlaIleGly 240

Qy 742 GCGAAGAGAGGTTTGGCGGAGGATCAGCGGTCTTCATCGACTTTTGTCTCAACCATGCC 801
Db 241 AlaGluGlyPheValSerThrSerIleAlaLysAsnTyLeuLeuAsnAla 260

Qy 802 AGAATATTATCTTTCAACCGCTATTCCGACCGCATGCTGTGCGGTCTTCACGAGCT 861
Db 261 ArgSerPheIlePheGlnThrAlaLeuSerProSerAlaIleGluAlaAlaArgGluGly 280

Qy 862 TTCAACATCATTTGAGCCAGCAGGAAACGACAGCTGTTTATTTCTTATCATCAGCATG 921
Db 281 IleSerIleIleGlnAsnGluProGluArgLysGlnLeuLysAsnAlaGlnTy 300

Qy 922 ATCAGAACCATGCTGGAAGAATATCGGTATGTGTGTAAGAGAGATCATCACACCATATT 981
Db 301 LeuArgLeuLysLeuGluSerGlyPheValMetLysGluGlyGluThrProIleIle 320

Qy 982 CCTGTAGTCAATGGCGATGCCATAAAACGCTCTATTGCTGAAAACTCGAGGGCAAG 1041
Db 321 SerLeuIleIleGlyGlySerHisGluAlaMetGlnPheSerAlaLysLeuLeuAspGlu 340

Qy 1042 GGAATTTATGCTCTGCTATTCGCGCCGCAACCGTTCCGCGCGTGAAGCCGATTCGA 1101
Db 341 GlyValPheIleProAlaIleArgProThrValProLysGlySerSerArgLeuArg 360
```

RESULT 8
G69647
Glycine C-acetyltransferase (EC 2.3.1.29) kbl - *Bacillus subtilis*
N;Alternate names: 2-amino-3-ketobutyrate CoA ligase kbl
C;Species: *Bacillus subtilis*
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: G69647
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Portetalle
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetalle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowaka, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: G69647
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-392 <N>
A;Cross-references: UNIPROT:O31777; UNIPARC:UPI0000060429; GB:Z99112; GB:AL009126; NID:9
A;Experimental source: strain 168
C;Genetics:
A;Gene: kbl
C;Superfamily: 5-aminolevulinatase synthase; glycine C-acetyltransferase homology
C;Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate
F;43-380/Domain: glycine C-acetyltransferase homology <GCA>
F;240/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:
Pred. No.: 1.56e-56 Length: 392
Score: 843.00 Matches: 169
Percent Similarity: 66.8% Conservative: 77
Best Local Similarity: 45.9% Mismatches: 120
Query Match: 15.3% Indels: 2
DB: 2 Gaps: 1


```
Qy 2805 GTGAATTTTTCATCAATTGATGCGCACCGGTTAGAGCGGTCAACGAATTAACCCG 2864
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 IleenPheLeuHisProIleLysGlyThrLysPheGlySerMetAspSerLeuThrPro 247
Qy 2865 CTGTATTGTTTAAAGTCTGCGCGTGTTCGGTGTTCATCAATCCATCAAAAGAAATTCGC 2924
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
248 MetLysCysLeuArgIleValAlaLeuPheArgLeuIleAsnProThrLysGluIleArg 267
Qy 2925 ATTTCGAGAGGAGAGGTCAATCTCCGACATTCGACATTCGACGATTCAGGCTTTAGCCGCA 2984
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 IleAlaGlyIleArgGluValAsnLeuArgSerLeuGlnProLeuAlaLeuLysAlaAla 287
Qy 2985 AACTCCATTTTTCGAGAGTACTTAACAATCGCGGCGCAAGAGAGACGAGGATCAT 3044
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
288 AsnSerIlePheValGlyAspTyrLeuIleThrGlyGlyGlnProAsnGlnLeuAspTyr 307
Qy 3045 AAAATGCTGAGTATTAGGCTTTGAAGTTGAA 3077
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
308 AspMetIleAsnAspLeuGlyPheGluIleAsp 318

RESULT 10
C84138
8-amino-7-oxononanoate synthase bioF [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Specie: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: C84138
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C84138
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-395 <STO>
A:Cross-references: UNIPROT:Q9K625; UNIPARC:UPI00000C4332; GB:AP0011520; GB:BA0000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: bioF
C:Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology
C:Keywords: phosphoprotein; pyridoxal phosphate
F:240/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Alignment Scores:
Pred. No.: 1,78e-46 Length: 395
Score: 712.00 Matches: 153
Percent Similarity: 62.5% Conservative: 74
Best Local Similarity: 42.1% Mismatches: 124
Query Match: 12.9% Indels: 12
DB: 2 Gaps: 5

US-10-681-086-1 (1-3156) x C84138 (1-395)
Qy 37 TGGTTAAAC-----GAGCGGTAGACAGAATGAAGAAGCGCGGTACATCGTAAAC 87
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5 TrpLeuHisAlaIleGluLysLeuThrArgLeuLysAspArgGlySerPheArgGln 24
Qy 88 CTGCGGTCAATGGATGGAGCGCGGTTCACAGAGGAATATGTATGGCGAAATCAACG 147
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
25 LeuValProThrSerGluAlaLeuProThrLeuThrArgGluAsnCysArgLeuLeu 44
Qy 148 GTCTGGTCTCAACAATATTATTAGCGCTCGCAAGCGATAGACGTTTTCATCGATGTCAGCC 207
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
45 AsnLeuAlaSerAsnAsnTyrLeuGlyIleAlaAspSerLysGluPheIleGluArgThr 64
Qy 208 CAACAGCATTTCAGCAATTTGGACAGGAAGACGCGGTTCACGTTTAAACGACAGGCAAT 267
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
65 GluGlnLeuAlaSerSerTyrAlaIleGlySerThrAlaSerArgLeuIleIleGlyAsn 84
Qy 268 TCGTCTGGCATGAAGACTAGAAAGAGATTGCCAGCTTTAACTGACAGAGACGCGCC 327
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
85 HisProLeuTyrGluGluAlaGluTyrGluLeuThrLysTyrLysLysThrGluAlaAla 104
Qy 328 CTGCTGTTTCGAGCGGTACTTGGCCCAATGTCGCTGTCTTTCATCTCCAGAGAAAG 387
```

```
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
105 LeuIlePheGlySerGlyTyrMetAlaAsnValGlyIleIleSerSerIleValGlyArg 124
Qy 388 GAAGATGCTATTTTAAGTACACAGCTCAATCATGCAAGTATGATGACGGCTCGGACTT 447
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 GlyAspAlaValPheSerAspLysLeuAsnHisAlaSerIleValAspGlyCysGlnLeu 144
Qy 448 TCTAAGCGCTACAGTTGTTTATCGGCATATTGATATGATGATCTTTGAAAACAAGCTG 507
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 SerArgAlaAspHisLeuArgPheArgHisAsnAspMetAspHisLeuGluThrLeuLeu 164
Qy 508 AATGAACACAGCGCTTATACGCGCGTTCCTTAACACAGCAGGAGTATTCAGCATGGAT 567
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165 GlnLysSerPro---HisLysGlnLysLeuIleValAlaAspAlaLeuPheSerMetAsp 183
Qy 568 GGCACAAATCGCCCTCTTGATCAGATCATCTCATCTCGGAAACGCTATCATCGCTTCGTG 627
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 GlyAspHisAlaAsnLeuHisAspLeuValThrLeuLysGluArgTyrGlyAlaIleLeu 203
Qy 628 GTCGTTGATGATGCCACGCAACAGGAGTTTTCGGCGATTTCGGGCAACAGGACGAGTGA 687
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 MetValAspGluAlaHisSerGlyValTyrGlyAlaThrGlyGlyLeuValGlu 223
Qy 688 TACTTTGCTGTGTTGCC-----GACATTGTTATCGGCACCTTAAGCAAAAGCTGTTGGC 741
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
224 GluLeuGlyLeuAsnAspArgValAspIleGlnMetGlyThrPheSerLysAlaLeuGly 243
Qy 742 GCGAAGAGAGGTTTTCGGCAGGATCAGCGGTCTTCATCGACTTTTTCGTAACCATGCC 801
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
244 SerTyrGlyTyrValAlaGlyAlaLysPheIleGluTyrLeuLeuAsnHisAla 263
Qy 802 AGAACATTTATCTTCAACCGCTATTCGCCACCGAGCTGTGCGGTGCTCAGGAGCT 861
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 ArgSerLeuIlePheThrThrAlaLeuProProTyrIleValAlaSerHisLeuAlaAla 283
Qy 862 TTCAACATCATTTGAAGCCAGC-----AGGAAAAACGACAGCTTTTATTTCTTTAT 912
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 LeuGlnIleValGlnGluGlnProTyrArgArgGluLysValGlnValLeuGly----- 301
Qy 913 ATCAGCATGATCAGAACCATGCTGAAGAANTATGGTTATGTGTGAAGGAGATCACACA 972
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 ---GluArgLeuArgAsnGlyLeuGluGlnLeuGlyPheSerLysCysArgLysSer 320
Qy 973 CCGATTATTCCTGTAGTCATTGGCGATGCCATAAAACGGTCTTATTTGCTGAAAAACTG 1032
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
321 TyrIleValProValLeuIleGlyAspAsnHisAspLeuLeuLeuValSerGluSerLeu 340
Qy 1033 CAGGCGCAAGGGAATTTATGCTCTCTCCATTCGCGCGCAACCGTTGCGCGGGTGAAGC 1092
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
341 GlnAlaAlaGlyIleAlaAlaIleProValArgProProThrValProArgGlyGluGly 360
Qy 1093 CGAATTGCA 1101
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 ArgIleArg 363

RESULT 11
A64462
8-amino-7-oxononanoate synthase (EC 2.3.1.47) - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: A64462
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
rsen, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: A64462
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-372 <BUL>
A:Cross-references: UNIPROT:Q58694; UNIPARC:UPI000012698C; GB:U67570; GB:L77117; NID:928
```


C;Genetics:

A;Map position: FOR1245358-1246476
C;Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology
F;32-363/Domain: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate
F;246/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Alignment Scores:

Pred. No.:	1.9e-45	Length:	372
Score:	698.50	Matches:	158
Percent Similarity:	62.6%	Conservative:	66
Best Local Similarity:	44.1%	Mismatches:	113
Query Match:	12.7%	Indels:	21
DB:	2	Gaps:	9

US-10-681-086-1 (1-3156) x A64462 (1-372)

```
Qy 40 TTAACAGCGGTTAGACAGATGAAGAGAGCGCGGTACATCGTAACCTGCGGTCAATG 99
Db 6 LeuArgArgGluIleGluIleLysAsnAsnGlyLeuTyrArgPheLeuArgLysLys 25
Qy 100 GATGAGCGCGGTTCCAGAGAGGAATATTGATGCGGAAATCAACCGGTCTGTCCTCA 159
Db 26 Asp-----AspGlyValLeuAep-----PheSerSer 34
Qy 160 AACAAATTATTAGGGCTCGCAAGCGATAGAGCTTTGATGATGCGGCCCAACAGCATTTG 219
Db 35 AsnAspTyrLeuCysLeuSerLysHisProGluValIleGluLysGluGlyLeu 54
Qy 220 CAGCAATTGGGACAGAGCGGTTTCAGTTTAAACGACAGCAATTCGGTCTGCAT 279
Db 55 ---LysTyrGlyAlaGlySerThrGlySerArgLeuThrSerGlyAsn---IleAsnHis 72
Qy 280 GAAAGCTAGAAAGAGATTGCCAGCTTTAACTGACAGAGCGGCTCTGTTTTCG 339
Db 73 GluArgLeuGluGluLysIleAlaGluPheLysGluThrGluArgThrLeuValTyrSer 92
Qy 340 AGCGGTTACTTGGCCATGTCGGTGTCTCTTTCATCTTGCAGAAAGAGAGATGTCATT 399
Db 93 SerGlyTyrAlaThrAsnValGlyValIleSerAlaLeuCysLysLysGlyAspLeuIle 112
Qy 400 TTAAGTACAGCGCTCAATCATGCAAGTATGATGACCGGCTGCGACATTTCAAGCGTAT 459
Db 113 LeuSerAspLysLeuAsnHisAlaSerIleIleAspGlyCysLysLeuSerLysAlaAsp 132
Qy 460 ACAGTTCTTATCGGCATATTGATGATGATGATGATGATGATGATGATGATGATGATG 516
Db 133 ValLeuIleTyrAsnHisCysAspValGluHisLeuThrAsnLeuIleGluGluAsnTyr 152
Qy 517 CAGCGTTATCAGCGCCCTTTTATCGTAACAGCGAGTATTACGATGATGCGCACAAATC 576
Db 153 GlyLysTyrAsnAsnLeuPheIleValThrAspGlyValPheSerMetAspGlyAspIle 172
Qy 577 GCCCCTCTGATCAGATCATCTCATCTTGGCAACCGCTATCATGCTTCTGTTGTTGAT 636
Db 173 AlaProLeuArgAspLeuLysLysIleAlaAspGluPheAsnAlaIleLeuIleAsp 192
Qy 637 GATGCCACGACAGAGTTTGGCGGATTCGGGACAGAGACGAGTGAATACTTTGGT 696
Db 193 AspAlaHisGlyThrGlyValLeuGlyAsp---GlyArgGlyThrLeuLysHisPheAsn 211
Qy 697 GTTGTCTCCGAC-----ATTGTTATCGGCACCTTAAGCAAGCTGTTTGGCGCGAA 747
Db 212 LeuLysProSerAspAsnIleValGlnIleGlyThrLeuSerLysAlaIleGlyGlyLeu 231
Qy 748 GAGGTTTTCGGCAGGATCAGCGGCTTTCATGCACTTTTGTGTAACCATGCGCAACA 807
Db 232 GlyGlyPheValCysGlyIleGluGluValValGluTyrIleLeuIleAsnThrSerArgSer 251
Qy 808 TTTATCTTTCAACCGCTATTCGCCAGCGAGCTGTCGGGCTGCTACAGGCTTTTCAAC 867
Db 252 PheIlePheSerThrAlaLeuProHisValValGluGlyCysIleLysAlaPheGlu 271
```

```
Qy 868 ATCAITTAAGCCAGCGGAAAAAACGACAGCTTTTATTTTATATCATCATCATCAGA 927
Db 272 ILeileGluLysThrAspIleValLysLysLeuGln---LysAsnIleLysIleAlaAsn 290
Qy 928 ACAGTCTGAAGATATGGGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 987
Db 291 LysValPheLysLysTyrGluPheIleLysGluAspAsnLeuThrProIleTyrProPhe 310
Qy 988 GTCATTTGGCGATGCCATATAAACCGTCTTATTTCTGAAAAAACTCGAGGCGCAAGGAAT 1047
Db 311 IlePheLysGlu-----LysThrMetGluIleAlaGluHisLeuIleLysAsnAsnIle 328
Qy 1048 TATGCTCTGCCATTCGCGCGCAACCGTTGCGCGCGGTGAAAGCCGATTCGA 1101
Db 329 PheCysValGlyIleArgTyrProThrValProLysGlyLeuGluAlaGlyIleArg 346
```

RESULT 12

F75017

probable glycine C-acetyltransferase (EC 2.3.1.29) PAB1244 - Pyrococcus abyssi (strain O
C;Species: Pyrococcus abyssi

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: F75017

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A;Reference number: A75001

A;Accession: F75017

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-398 <RAW>

A;Cross-references: UNIPROT:Q9UY32; UNIPARC:UPI00000345A5; GB:AJ248288; GB:AL096836; NID

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB1244

C;Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology
C;Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate
F;47-385/Domain: glycine C-acetyltransferase homology <GCA>
F;244/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:

Pred. No.:	6.1e-45	Length:	398
Score:	692.00	Matches:	151
Percent Similarity:	61.8%	Conservative:	74
Best Local Similarity:	41.5%	Mismatches:	133
Query Match:	12.5%	Indels:	6
DB:	2	Gaps:	4

US-10-681-086-1 (1-3156) x F75017 (1-398)

Qy 25 AAGATTGATTCTCTGGTTAAACGAGCGGTAGACAGAAATGAAGAAAGCGCGTACATCGT 84

Db 6 LysLeuAep---TrpIleLysGluGluLeuGluLysLysLysGlyLeuTyrVal 24

Qy 85 AACCTCGGTCAATGATGAGCGCGGTTCCAGAGAGGAATATTGATGCGGAAATCAA 144

Db 25 ThrIleArgValLeuGlnSerAlaGlnGlyProTyrValValValAsnGlyLysArgVal 44

Qy 145 ACGGTCTGGTCTCAACATATTATTAGGCTCGCAAGCGATAGACGTTTATCATGATGCA 204

Db 45 LeuAsnMetCysSerAsnAsnTyrLeuGlyLeuAlaHisProLysIleLysGluAla 64

Qy 205 GCCCAACACGATGTCAGCAATTTGGGACAGAGAGCGGTTCCACCGTTTCAACGACAGGC 264

Db 65 AlaIleArgAlaIleLeuAspTyrGlyValGlyAlaGlyAlaValArgThrIleAlaGly 84

Qy 265 AATTCGGTCTGGCATGAAAGCTAGAAAAGAGATTGCCAGCTTTAACTGACAGAGCG 324

Db 85 ThrMetGluLeuHisValGluLeuGluLysLeuAlaLysPheLysLysArgGluAla 104

Qy 325 GCCTCTGTTTTCGACCGGTACTTGGCCAAATGTCGGTCTCTTCATCTCTGCCAGAA 384

Db 105 AlaIleLeuPheGlnSerGlyTyrAsnAlaAsnLeuGlyAlaIleSerAlaLeuLeuArg 124

Alignment Scores:

```
Pred. No.: 1.76e-44 Length: 398
Score: 686.00 Matches: 150
Percent Similarity: 61.8% Conservative: 75
Best Local Similarity: 41.2% Mismatches: 133
Query Match: 12.4% Indels: 6
DB: 2 Gaps: 4

US-10-681-086-1 (1-3156) x E71454 (1-398)

QY 25 AAGATTGATCTCTGTTAAACAGCGGTTAGACAGAAATGAAGAAGCGCGGTACATCGT 84
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 LysLeuAsp---TrpIleYsgluGluLeuGluLeuLysLysLysLeuTyrVal 24

QY 85 AACCTGCGGTCAATGATGGAGCGCGGTTCCAGAGAGAAATATTGATGGCGAAATCAA 144
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 ThrIleArgValLeuGlnSerAlaGlnGlyProTrpIleValValAsnGlyLysArgVal 44

QY 145 ACGGTCTGCTCTCAAAACAATTTATTAGGCTCGCAAGCGATAGACGTTTGATCGATCA 204
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 LeuAsnMetCysSerAsnAsnTyrLeuGlyLeuAlaAlaHisProLysIleLysGluAla 64

QY 205 GCCCAACAGCATTCACCAATTTGGGACAGAGACGCGGTTCACGTTTACGACAGGC 264
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 AlaIleArgAlaIleLeuAspTyrGlyValGlyAlaGlyAlaValArgThrIleAlaGly 84

QY 265 AATTGCGTCTGCGATGAAAGCTAGAAAAGAGATTGCCAGCTTTAAACTGACAGAGCG 324
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 ThrMetGluLeuHisValGluLeuGluLysLeuAlaLysPheLysLysArgGluAla 104

QY 325 GCCTCTGCTGTTTCGAGCGGTTACTTGGCCAAATGTCGGTGTCTTTCATCTCTGCCAGAA 384
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 AlaIleLeuPheGlnSerGlyTyrAsnAlaAsnLeuGlyAlaIleSerAlaLeuLeuArg 124

QY 385 AAG-----GAAGATGTCATTTAAGTGCACAGCTCAATCATGCAAGTATGATCGACGC 438
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 LysGlyGluAspGlyValPheLeuSerGluLeuLeuAsnHisAlaSerIleIleAspGly 144

QY 439 TCGCGACTTCTAAGCGTGATACAGTGTGTTATCGGCATATTGATATGAATGATCTTGAA 498
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 MetArgLeuSerGlyAlaProLysValIleTyrLysHisLeuAspValAspAspLeuLys 164

QY 499 AACAGCTGAATGAACAACAGCGTTATCAGCGCGCGTTTATCGTAACACAGCGAGTATTC 558
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 165 LysLysLeuGluGluAsnLysAspLysLysLysIleIleValThrAspGlyValPhe 184

QY 559 AGCATGATGACACATCCGCTCTTGCATCAGATCATCTCCTTCGAAAACGCTATCAT 618
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 SerMetAspGlyAspLeuAlaProLeuProGluIleValGluValAlaGluGlnTyrAsp 204

QY 619 GCCTTCGTGTCGTTGATGCCACGACAGAGTTTGGGCGATTTCGGGACAGGA 678
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 AlaIleValTyrValAspAspAlaHisGlyGluGlyValLeuGlySerHisGlyArgGly 224

QY 679 ACGAGTGAATCTTTGTTGTTGTCCC-----GACATTGTTATCGGCACCTTAAGCAAA 732
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 225 IleValAspHisPheAsnLeuHisAspLysValAspPheGluMetGlyThrLeuSerLys 244

QY 733 GCTGTTGGCGGAGAGGTTTTCGGCAGGATCAGCGGTCTTCATCGACTTTTGTCTG 792
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 AlaPheGlyValIleGlyGlyTyrValAlaGlyProGluGluAlaIleGluTyrLeuLys 264

QY 793 AACCATGCCAGAACATTTATCTTCAACCGCTATTCGCCGACGAGTTCGGGCTGCT 852
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 265 GlnArgAlaArgProPheLeuPheSerSerAlaMetAsnProProAspValAlaAlaAla 284

QY 853 CACGAGGCTTTCAACATCAATGAACCCAGAGGAGGAAAAACGACGCTTTTATTTCTTAT 912
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 285 IleAlaAlaValGluIleLeuGlnLysSerAspLeuValLysLysLeuTyrAspAsn 304

QY 913 ATCAGCATGATCAGAACCGACTCTGAGNAATATGGTTATGTGTGTAAGAGGATCACACA 972
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 305 ThrHisPheGlnLysGlyLeuArgAspLeuGlyTyrAspLeuGlyAsnThrLysHis 324
```

Search completed: February 10, 2006, 03:00:23
Job time : 129 secs

```
QY 973 CCGATTATTCTCTAGTCAATTGGCGATGCCCAATAAACGGTCTCTATTGCTGAAAAAAGCTG 1032
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 325 ProIleThrProValMetLeuTyrAspGluLysLeuAlaGlnGluPheSerArgArgLeu 344

QY 1033 ---CAGGGCAAGGAATTTATGCTCTCCATTCCGCCGCGCAACCGTTGCGCGGCGTCAA 1089
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 345 TyrGluGluTyrAsnIlePheAlaGlnAlaIleValTyrProThrValProLeuGlyThr 364

QY 1090 AGCCGGATTGCA 1101
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 365 AlaArgIleArg 368
```

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

BEST AVAILABLE COPY

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2006, 02:17:36 ; Search time 139.9 Seconds

(without alignments)
3183.198 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggttacgacg.....atcggtatgatgtcgaattc 3156

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abses/ABSSWEB_spool/US10681086/runat_09022006_140515_11713/app_query.fasta_1
-DB=Uniprot -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUT=txt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USN=US10681086 @CGN 1.1 466 @runat_09022006_140515_11713 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -DBLOP=6 -DELEXT=7

Database : Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1836	33.3	389	1 B10F_BACSU	P53556 bacillus su
2	1676	30.4	371	2 Q8KZM9_BACSU	Q8KZM9 bacillus su
3	1350	24.5	256	2 Q7BU13_STAAU	Q7BU13 staphylococ
4	1350	24.5	256	2 Q6GKR2_STAAU	Q6GKR2 staphylococ
5	1350	24.5	256	2 Q87369_STAAU	Q87369 staphylococ
6	1350	24.5	256	2 Q7ABD0_STAAU	Q7ABD0 staphylococ
7	1333	24.2	253	1 KANU_STAAU	P05057 staphylococ
8	1333	24.2	253	2 Q7DU50_STAAU	Q7DU50 staphylococ
9	1327	24.1	253	1 KANU_BACSP	P05058 bacillus sp
10	1318	23.9	253	2 Q57514_BACST	Q57514 bacillus st
11	1317	23.9	335	1 B10B_BACSU	P53557 bacillus su
12	1308	23.7	335	2 Q8KZM7_BACSU	Q8KZM7 bacillus su
13	1286.5	23.3	379	2 Q65ML1_BACLD	Q65ML1 bacillus li
14	1222	22.1	253	2 Q75WE9_92Z2Z	Q75WE9 plasmid pub
15	1222	22.1	333	2 Q65MK9_BACLD	Q65MK9 bacillus li
16	1214	22.0	339	2 Q70JZ1_BACAM	Q70JZ1 bacillus am

17	1103	20.0	333	2	Q9KC26_BACHD	Q9KC26 bacillus ha
18	1050	19.0	333	2	Q5KZNI_GEOKA	Q5KZNI geobacillus
19	971	17.6	332	1	B10B_BACSH	P19206 bacillus sp
20	942	17.1	387	2	Q9AJN1_9BACL	Q9AJN1 kurtzia sp.
21	929	16.8	389	1	B10F_BACSH	P22806 bacillus sp
22	851	15.4	338	2	Q9AJN0_9BACL	Q9AJN0 kurtzia sp.
23	843	15.3	392	1	KBL_BACSU	O31777 bacillus su
24	828	15.0	313	2	Q8COB3_STABP	Q8COB3 staphylococ
25	828	15.0	321	2	Q5HKJ7_STABEQ	Q5HKJ7 staphylococ
26	809	14.7	332	2	Q4MMW5_BACEE	Q4MMW5 bacillus ce
27	809	14.7	332	2	Q6HE51_BACHK	Q6HE51 bacillus th
28	809	14.7	332	2	Q635G7_BACCZ	Q635G7 bacillus ce
29	809	14.7	332	2	Q81MB3_BACAN	Q81MB3 bacillus an
30	808	14.6	332	2	Q73112_BACC1	Q73112 bacillus ce
31	805	14.6	332	2	Q818X3_BACCR	Q818X3 bacillus ce
32	804	14.6	391	2	Q65JE6_BACLD	Q65JE6 bacillus li
33	793	14.4	335	2	Q6G6P6_STAAS	Q6G6P6 staphylococ
34	733	14.4	335	2	Q6GE08_STAAU	Q6GE08 staphylococ
35	793	14.4	335	2	Q5HDC9_STAAC	Q5HDC9 staphylococ
36	793	14.4	335	2	Q7A018_STAAW	Q7A018 staphylococ
37	793	14.4	335	2	Q7A3R9_STAAM	Q7A3R9 staphylococ
38	793	14.4	335	2	Q99RK7_STAAM	Q99RK7 staphylococ
39	778	14.1	321	2	Q4L9U7_STAHT	Q4L9U7 staphylococ
40	771	14.0	390	2	Q5KY23_GEOKA	Q5KY23 geobacillus
41	767	13.9	391	2	Q749W3_GEOSL	Q749W3 geobacter s
42	763.5	13.8	395	2	Q73KM3_TREDE	Q73KM3 treponema d
43	755	13.7	398	2	Q6KH66_MYCWO	Q6KH66 mycoplasma
44	743.5	13.5	400	2	Q5V3N9_HALMA	Q5V3N9 haloarcula
45	739.5	13.4	392	2	Q67N86_SYMTH	Q67N86 symbiobacte

ALIGNMENTS

RESULT 1
B10F_BACSU
ID B10F_BACSU STANDARD; PRT; 389 AA.
AC P53556;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE 8-amino-7-oxononanoate synthase (EC 2.3.1.47) (AONS) (8-amino-7-ketopelargonate synthase) (7-keto-8-amino-pelargonic acid synthetase)
DE (7-KAP synthetase) (L-alanine-piameyl COA ligase).
GN Name=biof; OrderedLocusNames=BSU30220;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96312354; PubMed=8763940;
RA Bower S., Perkins J.B., Yocum R.R., Howitt C.L., Rahaim P., Pero J.;
RT "Cloning, sequencing, and characterization of the Bacillus subtilis biotin biosynthetic operon."
RL J. Bacteriol. 178:4122-4130(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.-D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes in the 200 kb rmb-dnaB region."
RL Microbiology 143:3431-3441(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogawara N., Mosser I., Albertini A.M., Alloni G., Azevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruch C.V., Caldwell B., Capuano V., Carter N.M., Choi S.-K., Codani J.-J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Duesterhoef A., Ehrlich S.D., Emerson P.T., Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guileppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M.,
RA Klein C.-M., Joriss B., Karamata D., Kashiwara Y., Klaerr-Bianchard M.,
RA Kuno M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone P., Sekiguchi J., Sekowska A., Seror S.J., Serron P.,
RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E.,
RA Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*,";
RL Nature 390:249-256(1997).
CC -1- CATALYTIC ACTIVITY: 6-carboxyhexanoyl-CoA + L-alanine = 8-amino-7-
CC oxononanoate + CoA + CO(2).
CC -1- COPACTOR: Pyridoxal phosphate.
CC -1- PATHWAY: Cofactor biosynthesis; biotin biosynthesis; biotin from
CC 6-carboxyhexanoyl-CoA: step 1.
CC -1- SIMILARITY: Belongs to the class-II pyridoxal-phosphate-dependent
CC aminotransferase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U51868; AB17459.1; -; Genomic_DNA.
CC EMBL; AF008220; AAC00263.1; -; Genomic_DNA.
CC EMBL; Z99119; CAB15000.1; -; Genomic_DNA.
CC PIR; F69594; F69594.
CC HSSP; P07912; 1FC4.
CC Subtilist; BG11527; bioF.
CC InterPro; IPR004839; Aminotrans_I/II.
CC InterPro; IPR001917; Aminotrans_II.
CC InterPro; IPR004723; BioF.
CC Pfam; PF00155; Aminotran_1.2; 1.
CC TIGRFAMs; TIGR00858; bioF; 1.
CC PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
KW Biotin biosynthesis; Complete proteome; Pyridoxal phosphate;
KW Transferrase.
FT BINDING 237 237 Pyridoxal phosphate (covalent) (By
FT similarity).
SQ SEQUENCE 369 AA; 42581 MW; 02B31E16D68F7D6C CRC64;
Alignment Scores:
Pred. No.: 4.98e-131 Length: 389
Score: 1836.00 Matches: 359
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 33.3% Indels: 0
DB: 1 Gaps: 0
US-10-681-086-1 (1-3156) x BIOF_BACSU (1-389)

Db 21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
QY 142 CAAACGGTCTGGTCTCAACAATATTATTAGGGCTCCAGCGATAGACGTTTTCATCGAT 201
Db 41 GlnThrValTrpSerSerAsnAsnTyLeuGlyLeuAlaSerAspArgLeuIleAsp 60
QY 202 GCAGCCCAACAGCATTTCAGCAATTTGGGACAGAGCAAGCGGTTTCACGTTTAAACGACA 261
Db 61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
QY 262 GGCATTTTCGGTCTGGCATGAAAAGCTAGAAAAGAGATATGCCAGCTTTAACTGACAGAA 321
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100
QY 322 GCGGCCCTGCTGTTTCGAGCGGTACTTCGCCAATTCGGTTCCTTTCATCTTCCTGCCA 381
Db 101 AlaAlaLeuLeuPheSerSerGlyTyLeuAlaAsnValGlyValLeuSerSerLeuPro 120
QY 382 GAAAGGAAGATGTCATTTTAAGTGACAGCTCAATCATCATGATGATGATGATGATGATGAT 441
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
QY 442 CGACTTTTAAAGCTGATACAGTTGTTTATCGGCATATTCATATGATGATGATGATGATGAT 501
Db 141 ArgLeuSerLysAlaAspThrValValTyArgHisIleAspMetAsnAspLeuGluAsn 160
QY 502 AAGCTGATCAACACACAGCGTTATCAGCGCGTTTATCGTACAGACGGAGTATTCAGC 561
Db 161 LysLeuAsnGluThrGlnArgTyGlnArgPheIleValThrAspGlyValPheSer 180
QY 562 ATGGATGGCAATCGCCCTCTTGTGATCAGATCATCTCACTTGGCAACGCTATCATGCC 621
Db 181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTyHisAla 200
QY 622 TTCGTGGTCTTGATGATGCCACGCAACAGGAGTTTGGCGGATTCGGGCAACAGGAAAG 681
Db 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
QY 682 AGTGAATACCTTGGTCTGTCGCGCATGTTATCGGCACCTTAAGCAAGCTGTTGGC 741
Db 221 SerGluTyThrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
QY 742 GCGGAGGAGGTTTTCGGCGAGGATCAGCGGCTTTCATCGACTTTTTCCTGAACCATGCC 801
Db 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
QY 802 AGAATATTATCTTCAACCGCTATTCGCGCAGCGAGCTGTGCGGCTGCTCAGAGGCT 861
Db 261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaHisGluAla 280
QY 862 TTCACATCATCTGAAGCAGCAGCGGAAACGACAGCTTTTATTTCTTATATCAGCATG 921
Db 281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyIleSerMet 300
QY 922 ATCAGAACCATCTGAAGATATGGGTTATGTGTGAAGAGAGATCACACACCATATTATT 981
Db 301 IleArgThrSerLeuLysAsnMetGlyTyThrValValLysGlyAspHisThrProIleIle 320
QY 982 CCTGTAGTCATTCGCGATGCCATATAAACCGGTCCTATTTCCTGCTGAAAACTGACGGCAAG 1041
Db 321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
QY 1042 GGAATTTATGCTCTCGCATTCGGCGCCCAACCGTTTCGCGCGGTGAAACCGCGATTCGA 1101
Db 341 GlyIleTyAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 360
RESULT 2
Q8KZM9_BACSU
ID Q8KZM9_BACSU PRELIMINARY; PRT; 371 AA.
DC Q8KZM9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

8-amino-7-oxononanoate synthase.
 Name=bioF;
 GN Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 NP NUCLEOTIDE SEQUENCE.
 RA Sasaki M., Kawamura F., Kurusu Y.;
 RP "Genetic Analysis of an Incomplete bio Operon in a Biotin Auxotrophic
 RT Strain of Bacillus subtilis Natto OK2.";
 RL Biosci. Biotechnol. Biochem. 68:739-742(2004).
 DR EMBL; AB080666; BAC03241.1; -; Genomic_DNA.
 DR HSP; PF07912; IFC4.
 DR GO; GO:0008710; F:8-amino-7-oxononanoate synthase activity; IEA.
 DR CO; GO:0016769; F:transferase activity, transferring nitrogen. . .; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR GO; GO:0009102; P:biotin biosynthesis; IEA.
 DR InterPro; IPR004839; AminoTrans_I.
 DR InterPro; IPR001917; AminoTrans_II.
 DR InterPro; IPR004723; BioF.
 DR Pfam; PF00155; AminoTrans_1_2; 1.
 DR TIGRFAMs; TIGR00858; bioF_1.
 DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN 1.
 SQ SEQUENCE 371 AA; 40446 MW; FDF0204622D226F1 CRC64;
 Alignment Scores:
 Pred. No.: 7.9e-119 Length: 371
 Score: 1676.00 Matches: 330
 Percent Similarity: 93.3% Conservative: 6
 Best Local Similarity: 91.7% Mismatches: 6
 Query Match: 30.4% Indels: 18
 DB: 2 Gaps: 1
 US-10-681-086-1 (1-3156) x Q8K2M9_BACSU (1-371)
 QY 22 TTGAAGATTGATTCCTGGTTAAACGACGCGTTAGACAGAAATGAAGAAGCCGCGTACAT 81
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 1 MetLysileAspSerTrpLeuAsnAspArgLeuAspIleAlaLysGluAlaGlyValHis 20
 QY 82 CGTAACTCGGTCATGGATGGAGCCGCGTCCAGAGAGGAATATTGATGCGGAAAT 141
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 21 ArgAsnLeuArgSerMetAsnGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
 QY 142 CAAACGCTCTGGTCTCAACAATATTATTAGGCTCGCAAGCGATAGACGTTTGCATCAT 201
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgArgLeuIleAsp 60
 QY 202 GCAGCCCAACAGCATTTGCAGCAATTTGGGACAGGAAGCAGCGGTTCACTGTTAACGACA 261
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
 QY 262 GGCATTTCTGCTGGCATGAAGCTAGAAAGAGATTGCCAGCTTTTAACTGACAGAA 321
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysArgThrGlu 100
 QY 322 GCGCCCTCTGTTTTCGAGCGGTTACTTGGCCCAATGTCGGTGTCTTCTATCTTCCCA 381
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
 QY 382 GAAAGGAAGATGTCATTTTAAAGTACAGCTCAATCATGCAAGTATGATCGAGCGGTGC 441
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerIleIleAspGlyCys 140
 QY 442 CGACTTTCTAAGCTGTACAGTGTGTTTATCGCATATTCATGATGATGATGATGATGAAAC 501
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160
 QY 502 AAGCTGAATGAACACACAGCGGTTATCAGCGCGTGTATTCGTAACAGACGAGTATTTCAGC 561
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 161 LysLeuAsnGlnThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
 QY 562 ATGGATCGCAATTCGCCCTCTTTGATCAGATCATCTCATTTCGGAACACGCTATCATGCC 621
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 181 MetAspGlyThrIleValProLeuAspGlnIleIleSerLeuAlaLysArgTyrHisAla 200
 QY 622 TTCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 681
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyArgGlyThr 220
 QY 682 AGTGAATATCTTGTGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 741
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 221 SerGluTyrPheGlyValTyrProAspIleValIleGlyThrLeuSerLysAlaValGly 240
 QY 742 GCGGAGGAGGTTTTCGCGCAGGATCAGCGGCTCTTTCATGCTGCTGCTGCTGCTGCTGCTGCT 801
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 241 ThrGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
 QY 802 AGAACATTATCTTCAACCGCTATTTCGCGCAGCAGCTGTCGGTGTCTGCTGCTGCTGCTGCTGCT 861
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 261 ArgThrPheIlePheGlnThrAlaIleProProIleSerCysAlaAlaHisGluAla 280
 QY 862 TTCACATCATTTGAAGCCAGCGGAAAAACGACAGCTTTTATTTTCTTTATATCAGCATG 921
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 281 PheAsn----- 282
 QY 922 ATCAGAACCATGCTGAAGATATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 981
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 283 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 302
 QY 982 CCTGTAGTCAATGCGCATGCCATATAAACCCTGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 303 ProValValIleGlyAspAlaHisLysThrValIlePheAlaGluLysLeuGlnGlyLys 322
 QY 1042 GGAATTATGCTCTCTCCATTCGCGCCCAACCTGTCGCGCGGTGAAAGCCGGAATTCGA 1101
 Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 323 GlyIleTyrAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 342
 RESULT 3
 Q7BU13 STAAU PRELIMINARY; PRT; 256 AA.
 ID Q7BU13; STAAU PRELIMINARY; PRT; 256 AA.
 AC Q7BU13;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Aminoglycoside adenylyltransferase Aadd (Kanamycin resistance protein)
 DE (O-nucleotidyltransferase (4')).
 GN Name=aadd; Synonyms=ant (4');
 OS Staphylococcus aureus.
 OG Plasmid pSK41, and Plasmid pUB110.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC PLASMID=pSK41;
 RX MEDLINE=98389645; PubMed=9721269;
 RA Berg T., Firth N., Apisridej S., Hettiaratchi A., Leelaporn A.,
 RA Skurray R.A.;
 RT "Complete nucleotide sequence of pSK41: evolution of staphylococcal
 RL conjugative multiresistance plasmids.";
 RL J. Bacteriol. 180:4350-4359(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HUC19; PLASMID=pUB110;
 RX MEDLINE=20316808; PubMed=10858352;
 RA Oliveira D.C., Wu S.W., de Lencastre H.;
 RT "Genetic organization of the downstream region of the mecA element in
 RL methicillin-resistant Staphylococcus aureus isolates carrying
 different polymorphisms of this region.";
 RL Antimicrob. Agents Chemother. 44:1906-1910(2000).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=N315;
 RX MEDLINE=92183847; PubMed=1544435; DOI=10.1016/0014-5793(92)80039-J;
 RA Hiratsuki K., Asada K., Suzuki E., Okonogi K., Yokota T.;
 RT "Molecular cloning and nucleotide sequence determination of the

US-10-681-086-1 (1-3156) x Q6GKR2_STAAR (1-256)

QY 1570 ATGAGATAGTGAATGACCAATAATATGACTAGAGAGAGAGAGAGAGAGATGTTTCAT 1629

Db 1 MetArgileValAsnGlyProIleMetThrArgGluGluArgMetLysIleValHis 20

QY 1630 GAAATTAAGGAAACAAATATTGGATAATATGGGATGATGTTAAGGCTATTGGTGTATTAT 1689

Db 21 GluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyr 40

QY 1690 GGCTCTCTGGTCTGACATGATGGCCCTATTTCGGATATTGAGATGATGTTGTCATG 1749

Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMet 60

QY 1750 TCACAGAGAGAGAGAGAGTTCACCCATGAATGACACACCGGTGAGTGGAGGTGGAGTG 1809

Db 61 SerThrGluGluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluVal 80

QY 1810 AATTTTCATGACGAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCG 1869

Db 81 AspPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100

QY 1870 CTACACATGGTCAATTTTCTATTTCGCGATTTATGATTCAGTGGATCTTAGAG 1929

Db 101 LeuThrHisGlyGlnPheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120

QY 1930 AAGTGTATCAATCTGTAATCGGTAGAGCCCAACCGTTCACGATGCGATTGGCC 1989

Db 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140

QY 1990 CTTATCGTAGAGAGCTGTTTGAATATGACGAGCAATGGCGTAAATATTCGTGTGCAAGGA 2049

Db 141 LeuIleValGluLeuLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGly 160

QY 2050 CCGACAACTTCTACCATCTTTCAGTGTACAGGTAGCAATGGCAGGTGGCATGTTGATT 2109

Db 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180

QY 2110 GGCTGTGATCATGCTGTTATACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAG 2169

Db 181 GlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLys 200

QY 2170 CAATCAGATCTTCTCAGGTATTACCATCTGCGAGTTTCGTAATGCTCGTCAACTT 2229

Db 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220

QY 2230 TCCGACTCTGAGAACTTCTCGAATCGTAGAGAAATTTCTGGAATGGGATTCAGGAGTGG 2289

Db 221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240

QY 2290 ACAGAACGACGAGATATAGTGGATGTGTCAAAACGCATACCATTT 2337

Db 241 ThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 256

RESULT 5

O87369_STAAM PRELIMINARY; PRT; 256 AA.

AC O87369;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Kanamycin nucleotidyltransferase.

GN Name=adab; OrderedLocusNames=SAV0035;

OS Staphylococcus aureus (strain Mu50 / ATCC 700699).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=158878;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,

RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,

RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,

RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,

RA Hattori M., Ogasawara N., Hayaishi H., Hiramatsu K.,

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

RL Lancet 357:1225-1240 (2001).

DR EMBL; BA00017; BAB56197.1; -; Genomic_DNA.

DR PIR; S09565; S09565.

DR HSSP; P05057; 1KNY.

DR SNR; O87369; 4-256.

DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR InterPro; IPR02934; DNA_pol_beta.

DR Pfam; PF01909; NTP_transf_2; 1.

KW Complete proteome; Transferase.

SQ SEQUENCE 256 AA; 29166 MW; B45EF8CDBD00654E CRC64;

Alignment Scores:

Pred. No.: 5,35e-94 Length: 256

Score: 1350.00 Matches: 256

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 24.5% Indels: 0

DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x O87369_STAAM (1-256)

QY 1570 ATGAGATAGTGAATGACCAATAATATGACTAGAGAGAGAGAGAGATGTTTCAT 1629

Db 1 MetArgileValAsnGlyProIleMetThrArgGluGluArgMetLysIleValHis 20

QY 1630 GAAATTAAGGAAACAAATATTGGATAATATGGGATGATGTTAAGGCTATTGGTGTATTAT 1689

Db 21 GluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyr 40

QY 1690 GGCTCTCTGGTCTGACATGATGGCCCTATTTCGGATATTGAGATGATGTTGTCATG 1749

Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMet 60

QY 1750 TCACAGAGAGAGAGAGTTCACCCATGAATGACACACCGGTGAGTGGAGGTGGAGTG 1809

Db 61 SerThrGluGluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluVal 80

QY 1810 AATTTTCATGACGAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCG 1869

Db 81 AspPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100

QY 1870 CTACACATGGTCAATTTTCTATTTCGCGATTTATGATTCAGTGGATCTTAGAG 1929

Db 101 LeuThrHisGlyGlnPheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120

QY 1930 AAGTGTATCAATCTGTAATCGGTAGAGCCCAACCGTTCACGATGCGATTGGCC 1989

Db 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140

QY 1990 CTTATCGTAGAGAGCTGTTTGAATATGACGAGCAATGGCGTAAATATTCGTGTGCAAGGA 2049

Db 141 LeuIleValGluLeuLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGly 160

QY 2050 CCGACAACTTCTACCATCTTTCAGTGTACAGGTAGCAATGGCAGGTGGCATGTTGATT 2109

Db 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180

QY 2110 GGCTGTGATCATGCTGTTATACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAG 2169

Db 181 GlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLys 200

QY 2170 CAATCAGATCTTCTCAGGTATTACCATCTGCGAGTTTCGTAATGCTCGTCAACTT 2229

Db 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220

QY 2230 TCCGACTCTGAGAACTTCTCGAATCGTAGAGAAATTTCTGGAATGGGATTCAGGAGTGG 2289

```
|||||
101 LeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120
QY AAAGTGTATCAAACTGTAATAACGAGAGCCCAACGTTTCACGATGCGATTGTGCC 1989
|||||
121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140
QY CTATCGTAGAAGAGCTGTTTGAATATGACAGGCAATGGCGTAAATATTCGTGTCAGGA 2049
|||||
141 LeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGly 160
QY CCGACAAACATTTTACCATCTTGAAGTGTACAGGTAGTAGGCGGCGTAACTGAAGCAGTTAAG 2109
|||||
161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180
QY GGTCTGCATCATCGCATCTCTTATACAGCAGGCGCTCGGCTCTTAACCTGAAGCAGTTAAG 2169
|||||
181 GlyLeuHisHisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLys 200
QY CAATCAGATCTTCTTTCAGGTTATGACCATCTGTGCGCAGTTCGTAATGTCTGGTCAACTT 2229
|||||
201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220
QY TCCGACTCTGAGAACTTCTCGAATCGCTAGAGAAATTTCTGGAATGCGATTTCAGAGTGG 2289
|||||
221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240
QY ACAGAACGACGACGATATAGTGGATGTGTCAAAACGACATACCATTT 2337
|||||
241 ThrGluArgHisGlyIleValAppValSerLysArgIleProPhe 256
|||||
RESULT 7
KANU STAAU STANDARD; PRT; 253 AA.
ID _KANU STAAU STANDARD; PRT; 253 AA.
AC P05057;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Name=knt; Nucleotidyltransferase (EC 2.7.7.-) (Nec(R)).
GN Staphylococcus aureus.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP NUCLEOTIDE SEQUENCE
RX MEDLINE=86174350; PubMed=3007933; DOI=10.1007/BF00330534;
RA Mueller R.E., Ano T., Imanaka T., Aiba S.;
RT "Complete nucleotide sequences of Bacillus plasmids PUB110dB, PRBH1
and its copy mutants.";
RL Mol. Gen. Genet. 202:169-171(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-7.
RX MEDLINE=85006820; PubMed=6094428;
RA Matsumura M., Katakura Y., Imanaka T., Aiba S.;
RT "Enzymatic and nucleotide sequence studies of a kanamycin-inactivating
enzyme encoded by a plasmid from thermophilic bacilli in comparison
with that encoded by plasmid PUB110.";
RL J. Bacteriol. 160:413-420(1984).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86301853; PubMed=3744038;
RA Bashkurov V.I., Mil'china N.V., Prozorov A.A.;
RT "Nucleotide sequence and physical map of kanamycin-resistant plasmid
PUB110 from Staphylococcus aureus.";
RL Genetika 22:1081-1092(1986).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86206447; PubMed=3010356;
RA McKenzie T., Hoshino T., Tanaka T., Sueoka N.;
RT "The nucleotide sequence of PUB110: some salient features in relation
to replication and its regulation.";
RL Plasmid 15:93-103(1986).
RN [5]
```

```
|||||
221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240
QY ACAGAACGACGACGATATAGTGGATGTGTCAAAACGACATACCATTT 2337
|||||
241 ThrGluArgHisGlyIleValAspValSerLysArgIleProPhe 256
|||||
RESULT 6
Q7A8D0 STAAU PRELIMINARY; PRT; 256 AA.
ID Q7A8D0;
AC Q7A8D0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Name=knt; Nucleotidyltransferase
GN Name=adD; OrderedLocuNames=SA00033;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshino A.,
Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; BA000018; BAB41251.1; -; Genomic_DNA.
DR SMR; Q7A8D0; 4-256.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR002934; DNA pol beta.
DR Pfam; PF01909; NTP transf 2_1.
KW Complete proteome; Transferase.
SQ SEQUENCE 256 AA; 29166 MW; B45EF8CDBD00654E CRC64;
|||||
Alignment Scores:
Pred. No.: 5.35e-94 Length: 256
Score: 1350.00 Matches: 256
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 24.5% Indels: 0
DB: 2 Gaps: 0
US-10-681-086-1 (1-3156) x Q7A8D0 STAAU (1-256)
QY 1570 ATGAGATAGTGAATGGACCAATATATAGTACGAGAGAAAGATGATGTTTCAT 1629
Db 1 MetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHis 20
QY 1630 GAAATTAGGACGAATATGATAAATGCGATGATGTTAAGGCTATTGTTTAT 1689
Db 21 GluIleLysGluArgGlyLeuAspLysTyrGlyAspValLysAlaIleLysValTyr 40
QY 1690 GGCTCTCTGTGTCGTCAGACTGATGGCCCTATTTCGGATTTGAGATGATGTTGTCATG 1749
Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMet 60
QY 1750 TCAACAGAGGAGCAGAGATTGACCCATGAATGGACACCGGTGAGTGAAGGTGGAAGTG 1809
Db 61 SerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluVal 80
QY 1810 AATTTTGATAGCAGAGATTTCTACTAGATTATGCTCAGGTGGATTCAGATTGGCGCG 1869
Db 81 AsnPheAspSerGluGluLysLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100
QY 1870 CTTACACATGGTCAATTTTCTCTATTGTCGCGATTTATGATTCAGGTGGATCTAGAG 1929
|||||
```



```
DR Pfam; PF01909; NTP_transf_2; 1.
KW Antibiotic resistance; Direct protein sequencing; Plasmid;
SQ SEQUENCE 253 AA; 28825 MW; 9CA14603B2B5DC6 CRC64;
Alignment Scores:
Pred. No.: 3,03e-92 Length: 253
Score: 1327.00 Matches: 251
Percent Similarity: 99.6% Conservative: 1
Best Local Similarity: 99.2% Mismatches: 1
Query Match: 24.1% Indels: 0
DB: 1 Gaps: 0

US-10-681-086-1 (1-3156) x KANU_BACSP (1-253)
Qy 1579 GTGAATGACCAATAATAATGACTAGAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638
Db 1 MethanGlyProIleMetThrArgGluGluArgMetIleValHisGluLeuLys 20
Qy 1639 GAACGAATATGGATAAATATGGGATGATGTTAAGCTATTGCTGTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspIleTyrGlyAspValIleGlyValIleGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGCCCTATTCCGATATTCAGATGATGTTGTCATGTCACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60
Qy 1759 GAAGCAGAGATTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheAsp 80
Qy 1819 AGCGAAGAGATTCTACTAGATTATGATGATGATGATGATGATGATGATGATGATGAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
Qy 1879 GGTCAATTTTCTATTGTCGCAATTTATGATTCAGTGGATGATGATGATGATGATGAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTTCGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaIleSerValGluAlaGlnIlePheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTGTAATATGAGGCAAAATGCGTAATATTCGTGTGCAAGGACCGACAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTCTACATCTCTGACTGTACAGGTAGCAATGCGAGTGCATGCGATGTTGATGTTGCTGCT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATGCGATCTGTTATACGAGCGCGCTTCGTTACTTAACGAGTGAAGCAATCAGAT 2178
Db 181 HisArgIleCysIleThrThrSerAlaSerValLeuThrGluAlaValIleGlnSerAsp 200
Qy 2179 CTTCCTTCAGGTTATGACCATCTGTGCGCAGTTCTGAATGCTGTGTCACACTTCCTGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTTCGGAATCGCTAGAGAAATTCGGAATCGGATTCAGAGTGCAGACGACGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy 2299 CACGGATATATAGTGGATGTTCAAAACGATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 10
ID Q57514_BACST PRELIMINARY; PRT; 253 AA.
AC Q57514;1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Kanamycin nucleotidyltransferase.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Narumi I., Nakayama N., Nakamoto S., Kihara H.;
RT "Complete nucleotide sequence of pSTK1, a cryptic plasmid from
RL Bacillus stearothermophilus TK015.";
RL Biotechnol. Lett. 15:1013-1016(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Narumi I., Nakayama N., Nakamoto S., Kihara H.;
RT "Bacillus stearothermophilus plasmid pSTK1 replicon is functional in
RL Escherichia coli.";
RL Biotechnol. Lett. 17:475-480(1995).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86120983; PubMed=3003740;
RA Liao H., McKenzie T., Hageman R.;
RT "Isolation of a thermostable enzyme variant by cloning and selection
RL in a thermophile.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:576-580(1986).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Narumi I., Nakayama N., Nakamoto S., Kimura T., Yanagkisawa T.,
RA Kihara H.;
RT "Construction of a new shuttle vector pSTE33 and its stabilities in
RL Bacillus stearothermophilus, Bacillus subtilis, and Escherichia
RL coli.";
RL Biotechnol. Lett. 15:815-820(1993).
DR EMBL; D29979; BAA06249.1; -; Genomic DNA.
DR EMBL; D29978; BAA06245.1; -; Genomic DNA.
DR HSSP; P05057; 1KNY.
DR SMR; Q57514; 1-253.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR002934; DNA_pol_beta.
DR Pfam; PF01909; NTP_transf_2; 1.
DR XTransferase.
SQ SEQUENCE 253 AA; 28873 MW; FE573F27646F281E CRC64;

Alignment Scores:
Pred. No.: 1,47e-91 Length: 253
Score: 1318.00 Matches: 250
Percent Similarity: 99.2% Conservative: 1
Best Local Similarity: 98.8% Mismatches: 2
Query Match: 23.9% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x Q57514_BACST (1-253)
Qy 1579 GTGAATGACCAATAATAATGACTAGAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetIleValHisGluLeuLys 20
Qy 1639 GAACGAATATGGATAAATATGGGATGATGTTAAGCTATTGCTGTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspIleTyrGlyAspValIleGlyValIleGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGCCCTATTCCGATATTCAGATGATGTTGTCATGTCACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60
Qy 1759 GAAGCAGAGATTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheTyr 80
Qy 1819 AGCGAAGAGATTCTACTAGATTATGATGATGATGATGATGATGATGATGATGATGAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
```

```
Qy 1879 GGTCAATTTTCTCTATTTCCTGATTTATGATTCAGGTGGATGATCTAGAGAAAGTGTAT 1938
Db |||||PhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAATCTGTAATCGGTAGAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTA 1998
Db |||||AlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTTGAATATGACGAGGCAATATGCGGTATATTCGTGTGCAAGGACCCACACA 2058
Db |||||GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGTACGTGTACAGTGTAGCAGTGTGCGATGTCATGTTGATTTGCTGTGAT 2118
Db |||||PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGCATCTCTTATACGACGAGCGCTTCGGTCTTAATCTGAAGCAGTTTAAGCAATCAGAT 2178
Db |||||HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
Qy 2179 CTTCTCTTCAGGTTATGACCATCTGCGCAGTTCGTATATGCTGCTCACTTTCGCACTCT 2238
Db |||||LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTTCTGGAATCGTACAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA 2298
Db |||||GluLysLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy 2299 CACGGATATATAGTGGATGTGTCAAAACGCAATACCATTT 2337
Db |||||HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

RESULT 11

```
BIOS_BACSU STANDARD; PRT; 335 AA.
AC P53557;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase).
GN Name=biob; OrderedLocusNames=BSU30200;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96312354; PubMed=8763940;
RA Bower S., Perkins J.B., Yocum R.R., Howitt C.L., Rahaim P., Pero J.;
RT "Cloning, sequencing, and characterization of the Bacillus subtilis
RT biotin biosynthetic operon."
RL J. Bacteriol. 178:4122-4130(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.-D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb rnb-dnaB region."
RL Microbiology 143:3431-3441(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.-K., Codani J.-C., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Duescherhoef A., Enrich S.D., Emmerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
```

```
RA Hilbert H., Holsappel S., Hosono S., Hullo M.-F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Konigstein G., Krogh S.,
RA Kumano M., Kurica K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivoita C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Takakoshi A., Tanaka T., Terptrata P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Medler H.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.
CC -1- COFACTOR: Binds 1 4Fe-4S cluster coordinated with 3 cysteines and
CC an exchangeable S-adenosyl-L-methionine (By similarity).
CC -1- COFACTOR: Binds 1 2Fe-2S cluster coordinated with 3 cysteines and
CC 1 arginine (By similarity).
CC -1- PATHWAY: Cofactor biosynthesis; biotin biosynthesis; biotin from
CC 6-carboxyhexanoyl-CoA: step 4 [final step].
CC -1- SIMILARITY: Belongs to the biotin and lipoic acid synthetases
CC family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; U51868; AB17461.1; -; Genomic DNA.
DR EMBL; AF008220; AAC00265.1; -; Genomic DNA.
DR EMBL; Z99119; CAB14998.1; -; Genomic DNA.
DR PIR; D69594; D69594.
DR Subtilist; BG11525; biob.
DR InterPro; IPR010722; BATS.
DR InterPro; IPR002684; Biotin synth.
DR InterPro; IPR006638; Elp3/Miab/Niib.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF06968; BATS; 1.
DR PIRSF; PIRSF001619; Biotin synth; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00433; biob; 1.
DR 2Fe-2S; 4Fe-4S; Biotin biosynthesis; Complete proteome; Iron;
KW Iron-sulfur; Metal-binding; Transference.
FT METAL 65 65 Iron-sulfur 1 (4Fe-4S-S-AdoMet) (By
FT similarity).
FT METAL 69 69 Iron-sulfur 1 (4Fe-4S-S-AdoMet) (By
FT similarity).
FT METAL 72 72 Iron-sulfur 1 (4Fe-4S-S-AdoMet) (By
FT similarity).
FT METAL 109 109 Iron-sulfur 2 (2Fe-2S) (By similarity).
FT METAL 141 141 Iron-sulfur 2 (2Fe-2S) (By similarity).
FT METAL 201 201 Iron-sulfur 2 (2Fe-2S) (By similarity).
FT METAL 271 271 Iron-sulfur 2 (2Fe-2S) (By similarity).
SQ SEQUENCE 335 AA; 36927 MW; D4B2E1A53271ED36 CRC64;
Alignment Scores:
Pred. No.: 1 87e-91 Length: 335
Score: 1337.00 Matches: 263
Percent Similarity: 92.0% Conservativeness: 2
Best Local Similarity: 91.3% Mismatches: 5
Query Match: 23.9% Indels: 18
DB: 1 Gaps: 1
```

US-10-681-086-1 (1-3156) x BIOB_BACSU (1-335)

QY 2307 TATAGTGGATGTGCAAAACCCATACCATTTTGAAT----- 2342
Db 48 TyrGlyLysLysValLysLeuAsnMetIleMetAsnAlaLysSerGlyLeuCysProGlu 67
QY 2343 -----TCGAAAGCGCGGATTTGATCTTACCGGATG 2372
Db 68 AsnCysGlyTyrCysSerGlnSerAlaIleSerLysAlaProIleGluSerTyrArgMet 87
QY 2373 GTGATAAGGAAACGCTCTTGAGGCGGAGCGGCGGACGATCTGAATATCGGCACA 2432
Db 88 ValAsnLysGluThrLeuLeuGluGlyAlaLysArgAlaHisAspLeuAsnIleGlyThr 107
QY 2433 TATTGTATCGTGGCAAGCGGACAGAGTCCGCTTACAGAGAAGTGCATCGTGTAGAT 2492
Db 108 TyrCysIleValAlaSerGlyArgGlyProSerAsnArgGluValAspGlnValValAsp 127
QY 2493 GCGGTTCCAGAAATTAAGAGAGCGGACGATGATGAGATTTTGGCATGTCTTGGACTGTG 2552
Db 128 AlaValGlnGluIleLysGluThrTyrGlyLeuLysValCysAlaCysLeuGlyLeuLeu 147
QY 2553 AAGCCAGACGAGCGGAGCGGCTCAAAGATGACGAGTACCGCTATATCAATAATTG 2612
Db 148 LysProGlnGlnAlaLysArgLeuLysAspAlaGlyValAspArgTyrAsnHisAsnLeu 167
QY 2613 AATACGTCACAGAGAAACCATTCAAACATCACACCTCACATACATACAGATGACAGATC 2672
Db 168 AsnThrSerGlnArgAsnHisSerAsnIleThrThrSerHisThrTyrAspAspArgVal 187
QY 2673 AATACGTTGAATCGCAAAAGAAATCGGGGCTGTCTCCGTGTTCCAGCGCCCATTCGGG 2732
Db 188 AsnThrValGluIleAlaLysGluSerGlyLeuSerProCysSerGlyAlaIleIleGly 207
QY 2733 ATGAAGGAGAGCAACAGGATGTGATGACATCGCCAAAAGCTTGAAGCTCTTGGACGCG 2792
Db 208 MetLysGluThrLysGlnAspValIleAspIleAlaLysSerLeuLysAlaLeuAspAla 227
QY 2793 GATTCCTTCGCTGGAATTTTTCATGCAATTCATGCGACGCTGTTAGAGCGGTCAAC 2852
Db 228 AspSerIleProValAsnPheLeuHisAlaIleAspGlyThrProLeuGluGlyValAsn 247
QY 2853 GAATTAACCCGCTGTATTGTTTAAAGTGTGGCGTGTTCGCTTTTATCAATCCATCA 2912
Db 248 GluLeuAsnProLeuTyrCysLeuLysValLeuAlaLeuPheArgPheIleAsnProSer 267
QY 2913 AAAGAAATTCGATTCGAGAGAGAGAGTCAATCTCCGCAATTGCGACCCATTAGGG 2972
Db 268 LysGluIleArgIleSerGlyArgGluValAsnLeuArgThrLeuGlnProLeuGly 287
QY 2973 CTTTACGCGCAAACTCCATTTTGTGCGAGACTACTTAACTGCGCGGCGCAGAGGAG 3032
Db 288 LeuTyrAlaAlaAsnSerIlePheValGlyAspTyrLeuThrThrAlaGlyGlnGluGlu 307
QY 3033 ACGGAGATCATAAATGCTCAGTGAATTTAGGCTTTGAAGTTGAATCAGTCCGAGAAATG 3092
Db 308 ThrGluAspHisLysMetLeuSerAspLeuGlyPheGluValGluSerValGluGluMet 327
QY 3093 AAGCGTGTGTTAAGTCGAAAGC 3116
Db 328 LysAlaSerLeuSerAlaLysSer 335

RESULT 12
Q8KZM7_BACSU PRELIMINARY; PRT; 335 AA.
AC Q8KZM7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Biotin synthetase.
GN Name=biob;
OS Bacillus subtilis.

Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=1423;
[1]
NUCLEOTIDE SEQUENCE.
Saaki M., Kawamura F., Kuruu Y.;
"Genetic Analysis of an incomplete bio Operon in a Biotin Auxotrophic
Strain of Bacillus subtilis Natto OK2.";
Biosci. Biotechnol. Biochem. 68:739-742(2004).
-I- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.
-I- COFACTOR: Binds a 4Fe-4S cluster coordinated with 3 cysteines and
an exchangeable S-adenosyl-L-methionine, and a 2Fe-2S cluster
coordinated with 3 cysteines and 1 arginine (by similarity).
-I- PATHWAY: Biotin biosynthesis; last step.
EMBL; AB088066; BAC03243.1; -; Genomic_DNA.
HSP; P12996; IR30.
GO; GO:0004076; P:biotin synthase activity; IEA.
GO; GO:0005506; P:iron ion binding; IEA.
GO; GO:0016740; P:transferase activity; IEA.
GO; GO:0009102; P:biotin biosynthesis; IEA.
DR InterPro; IPR010722; BATS.
DR InterPro; IPR002684; Biotin synth.
DR InterPro; IPR006638; Elp3/MiAB/NifB.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF06968; BATS; 1.
DR Pfam; PF04055; Radical SAM; 1.
DR PIRSF; PIRSF001619; Biotin synth; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00433; bioB; 1.
KW 2Fe-2S; 4Fe-4S; Biotin biosynthesis; Iron; Iron-sulfur; Metal-binding;
KW Transferase.
SQ SEQUENCE 335 AA; 36980 MW; 82693C6E168D9FAC CRC64;

Alignment Scores:
Pred. No.: 9 08e-91 Length: 335
Score: 1308.00 Matches: 260
Percent Similarity: 92.0% Conservative: 5
Best Local Similarity: 90.3% Mismatches: 5
Query Match: 23.7% Indels: 18
DB: Gaps: 1

US-10-681-086-1 (1-3156) x Q8KZM7_BACSU (1-335)

QY 2307 TATAGTGGATGTGCAAAACCCATACCATTTTGAAT----- 2342
Db 48 TyrGlyLysLysValLysLeuAsnMetIleMetAsnAlaLysSerGlyLeuCysProGlu 67
QY 2343 -----TCGAAAGCGGATTTGATCTTACCGGATG 2372
Db 68 AsnCysGlyTyrCysSerGlnSerAlaIleSerLysAlaProIleGluSerTyrArgMet 87
QY 2373 GTGATAAGGAAACGCTCTTGAGGCGGAGCGGCGGACGATCTGAATATCGGCACA 2432
Db 88 ValAsnLysGluThrLeuLeuGluGlyAlaLysArgAlaHisAspLeuAsnIleGlyThr 107
QY 2433 TATTGTATCGTGGCAAGCGGACAGAGTCCGCTTACAGAGAAGTGCATCGTGTAGAT 2492
Db 108 TyrCysIleValAlaSerGlyArgGlyProSerAsnArgGluValAspGlnValValAsp 127
QY 2493 GCGGTTCCAGAAATTAAGAGAGCGGACGATGATGAGATTTTGGCATGTCTTGGACTGTG 2552
Db 128 AlaValGlnGluIleLysGluThrTyrGlyLeuLysValCysAlaCysLeuGlyLeuLeu 147
QY 2553 AAGCCAGACGAGCGGAGCGGCTCAAAGATGACGAGTACCGCTATATCAATAATTG 2612
Db 148 LysProGlnGlnAlaLysArgLeuLysAspAlaGlyValAspArgTyrAsnHisAsnLeu 167
QY 2613 AATACGTCACAGAGAAACCATTCAAACATCACACCTCACATACATACAGATGACAGATC 2672
Db 168 AsnThrSerGlnArgAsnHisSerAsnIleThrThrSerHisThrTyrAspAspArgVal 187
QY 2673 AATACGTTGAATCGCAAAAGAAATCGGGGCTGTCTCCGTGTTCCAGCGCCCATTCGGG 2732
Db 188 AsnThrValGluIleAlaLysGluSerGlyLeuSerProCysSerGlyAlaIleIleGly 207


```
QY 922 ATCAGAACCACTGTAAGAATATGGTATATGTGTGAAGAGAGATCAACACCAATATT 981
Db 922 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 922 CCGTGTAGTGGGATGCCATTAACACGCTCTATTGCTGTAACAACTGCAGGGCAAG 1041
Db 922 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 922 ProValMetIleGlyAspProGlnLysAlaValGlnPheAlaAenGlyLeuLysGluLys 329
Db 922 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1042 GGAATTTATGCTCTGCGCATTCGCGCCGCAACCGTTGCGCGGTGAAGCCGATTCGA 1101
Db 1042 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 GlyIleGluAlaProAlaIleArgProProThrValAlaGluGlyGluSerArgIleArg 349
Db 330 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q75WE9_922ZZ PRELIMINARY; PRT; 253 AA.
AC Q75WE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Highly thermostable kanamycin nucleotidyltransferase
DE (EC 2.7.7.46).
GN Name=htk;
OS Plasmid PUB110.
OC other sequences; plasmids.
OX NCBI_TaxID=2655;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20013005; PubMed=10544290;
RA Haseki J., Yano T., Koyama Y., Kuramitsu S., Kagamiyama H.;
RT "Directed evolution of thermostable kanamycin-resistance gene: a
RT convenient selection marker for Thermus thermophilus.";
RL J. Biochem. 126:951-956(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21483306; PubMed=11602251; DOI=10.1016/S0014-5793(01)02926-X;
RA Hashimoto Y., Yano T., Kuramitsu S., Kagamiyama H.;
RT "Disruption of Thermus thermophilus genes by homologous recombination
RT using a thermostable kanamycin-resistant marker.";
RL FEBS Lett. 506:231-234(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Haseki J., Yano T., Koyama Y., Kuramitsu S., Kagamiyama H.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB121443; BAC98411.1; -; Genomic_DNA.
DR SMR; Q75WE9; 1-253.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR002934; DNA_pol_beta.
DR Pfam; PF01909; NTP_transf_2; 1.
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 253 AA; 28860 MW; BC691E32B8E5705E CRC64;

Alignment Scores:
Pred. NO.: 3,1e-84 Length: 253
Score: 1222.00 Matches: 231
Percent Similarity: 93.7% Conservative: 6
Best Local Similarity: 91.3% Mismatches: 16
Query Match: 22.1% Indels: 0
DB: Gaps: 0

US-10-681-086-1 (1-3156) x Q75WE9_922ZZ (1-253)

QY 1579 GTGAATGACCAATAAATGACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1638
Db 1579 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1639 GAACGAATATGGAATAATATGCGGATGATGTTAAGCGCTATGTTGTTTATGGCTCTCT 1698
Db 1639 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 21 GluArgIleLeuAspLysTyGlyAspValLysAlaIleGlyValTyGlySerLeu 40
Db 21 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1699 GGTGCTGAGATGATGGGCGCTATTCGAGATTCGAGATGATGTTGCTGATCAACAGAG 1758
Db 1699 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 41 GlyArgGlnThrAspGlyProTySerAspIleGluMetMetCysValLeuSerThrGlu 60
Db 41 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 1759 GAAGCAGAGTTCCAGCCATGAATGACCAACCGTGTGAGTGGAGGTGGAAGTGAATTTTGTAT 1818
Db 1759 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1819 AGCGAAGAGATTTCTACTAGATATGATCTCAGGTGGAATCAGATTTGGCCGCTTACACAT 1878
Db 1819 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1879 GGTCAATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGATCTTATAGAAAAGTGTAT 1938
Db 1879 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1939 CAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATCGGATTTGTGCCCTTATCGTA 1998
Db 1939 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1999 GAAGAGCTGTTGATATGACGAGCAATGCGGTATATTCGTGTGCAAGGACCCACACA 2058
Db 1999 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2059 TTTCTACCATCTTGACTGTACAGTAGCAATGCGAGTGCATGTTGATTGGTCTGCAT 2118
Db 2059 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2119 CATCGCATCTGTTATACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGAT 2178
Db 2119 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2179 CTTCTTCAGGTTATGACCATCTGTGCGAGTTCGATGCTGATGCTGCTCACTTCGAGCTCT 2238
Db 2179 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2239 GAGAACTCTCGAATCGTAGAGAAATTTCTGGAATGCGATTCAGGAGTGCACAGAACA 2298
Db 2239 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2299 CACGATATATAGTGGATGTGTCAAACGCAATACCATTT 2337
Db 2299 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
Db 241 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
Q65MK9_BACLD PRELIMINARY; PRT; 333 AA.
AC Q65MK9; Q62Y02;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE BioB (Biotin synthetase).
GN Name=biob; OrderedLocusNames=BL00956, BL100770;
OS Bacillus licheniformis (strain DSM 13 / ATCC 14580).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15383718; DOI=10.1159/000079829;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baeumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The complete genome sequence of Bacillus licheniformis DSM13, an
RT organism with great industrial potential.";
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., Lopez de Leon A., Xiang H., Gusti V., Clausen I.G.,
RA Olsen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.L.,
RA Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N.,
RA Erlich S.D., Berk R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2006, 02:56:51 ; Search time 19 Seconds
(without alignments)

2746.573 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggttacgagc.....atcggtatgatgcgaattc 3156

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 114120

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ARSSWEB_spool/US10681086/runat_09022006_140520_11815/app_query.fasta_1
-DB=issued Patents AA -QMT=fastan -SUFFIX=n2p.ral -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USPR=US10681086 -CGN 1 1 71 @runat_09022006_140520_11815
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : Issued Patents AA:

1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/PCUS COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1333	24.2	253	1	US-08-642-045B-4
2	1333	24.2	253	2	Sequence 4, Appli
3	1333	24.2	253	2	Sequence 2, Appli
4	1318	23.9	253	2	Sequence 11, Appl
5	1282	23.2	253	2	Sequence 1, Appli
6	1277	23.1	253	2	Sequence 12, Appl
7	1268	23.0	253	2	Sequence 14, Appl
8	1267	23.0	253	2	Sequence 15, Appl
9	1267	23.0	253	2	Sequence 2, Appli
10	1265	22.9	253	2	Sequence 13, Appl
11	1262	22.9	253	2	Sequence 20, Appl
12	1262	22.9	253	2	Sequence 17, Appl
					Sequence 18, Appl

13	1260	22.8	253	2	US-09-697-186B-16
14	1256	22.8	253	2	Sequence 16, Appl
15	1222	22.1	253	2	Sequence 19, Appl
16	942	17.1	387	2	Sequence 3, Appli
17	942	17.1	387	2	Sequence 6, Appli
18	942	17.1	387	2	Sequence 6, Appli
19	942	17.1	387	2	Sequence 6, Appli
20	851	15.4	338	2	Sequence 8, Appli
21	851	15.4	338	2	Sequence 8, Appli
22	851	15.4	338	2	Sequence 8, Appli
23	851	15.4	338	2	Sequence 8, Appli
24	828	15.0	327	2	Sequence 8, Appli
25	828	15.0	327	2	Sequence 3471, Ap
26	828	15.0	332	2	Sequence 540, App
27	704	12.8	414	2	Sequence 2562, Ap
28	673.5	12.2	398	2	Sequence 5771, Ap
29	673.5	12.2	398	2	Sequence 12, Appl
30	673.5	12.2	398	2	Sequence 12, Appl
31	673.5	12.2	398	2	Sequence 12, Appl
32	660	12.0	512	2	Sequence 12, Appl
33	612.5	11.1	488	2	Sequence 13924, A
34	602.5	10.9	387	2	Sequence 11131, A
35	602.5	10.9	387	2	Sequence 2, Appli
36	602.5	10.9	387	2	Sequence 6, Appli
37	591.5	10.7	412	2	Sequence 8, Appli
38	590.5	10.7	436	2	Sequence 10771, A
39	587.5	10.6	411	2	Sequence 19792, A
40	567.5	10.3	771	2	Sequence 7404, Ap
41	559	10.1	391	2	Sequence 157, App
42	546.5	9.9	394	2	Sequence 5575, Ap
43	535	9.7	384	1	Sequence 4736, Ap
44	535	9.7	384	1	Sequence 2, Appli
45	535	9.7	384	2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-08-642-045B-4

; Sequence 4, Application US/08642045B

; Patent No. 5851804

; GENERAL INFORMATION:

; APPLICANT: Snyder, Linda A.

; APPLICANT: Satishchandra, C.

; TITLE OF INVENTION: CHIMERIC KANAMYCIN RESISTANCE GENE

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5851804ris

; STREET: One Liberty Place, 46th floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: WINDOWS

; SOFTWARE: WordPerfect 6.0/6.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/642,045B

; FILING DATE: 06-MAY-1996

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: DeLuca, Mark

; REGISTRATION NUMBER: 33,229

; REFERENCE/DOCKET NUMBER: APOL-0262

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-568-3100

; TELEFAX: 215-568-3439

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 253 amino acids

; TYPE: amino acid

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-045B-4

Alignment Scores:
Pred. No.: 6.5e-138 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 1 Gaps: 0

US-10-681-086-1 (1-3156) x US-08-642-045B-4 (1-253)

QY 1579 GTGAATGACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20

QY 1639 GAACGAATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTCTCAGACTGATGGCCCTATTCCGATATTGAGATGATGTGTTCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60

QY 1759 GAACGACAGTTCCACCATGAATGACCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGCAT 1818
Db 61 GluAlaGluPheSerHisGluIleThrThrGlyGluIleValGluValAsnPheAsp 80

QY 1819 AGCGAAGAGATTCTACTAGATTATGATATTGAGATGATGTGTTCATGTCAACAGAG 1878
Db 81 SerGluGluIleLeuAspTyrAlaSerGlnValGluSerAspTyrProLeuThrHis 100

QY 1879 GGTCAATTTTCTCTATTTGCCGATTTATGATTTCAGTGGAAATCAGATTGCCGCTTACAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120

QY 1939 CAAACTGCTAATCGGTAGACCCCAACCGTCCACGATGGATTGTGGCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140

QY 1999 GAAGAGCTGTTTGAATATGACGCAAAATGGCGTAAATATTCGTGTGCAAGGACCGCAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrPargAsnIleArgValGlnGlyProThrThr 160

QY 2059 TTTCTACCATCTTGAATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

QY 2119 CATCGCATCTGTATACGACGACGCTTCGGTCTTAATGAGTGAAGTGAAGTGAAGTGAAG 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200

QY 2179 CTTCTCTCAGGTTATGACCATCTGTCGACGATTCGTAATGCTCGTCAACTTTCCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

QY 2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACGACGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheThrAsnGlyIleGlnGluIleThrThrGluArg 240

QY 2299 CACGATATATGATGATGCTGTCAAACGCCATACCATT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 2
US-08-852-268-2
; Sequence 2, Application US/08852268
; Patent No. 6143527
; GENERAL INFORMATION:
; APPLICANT: Pachuk, Catherine J.
; APPLICANT: Samuel, Manoj
```

```
QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACCGTTCACGATGCGATTTGTGCCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGAATATGACGAGCAATGCGGTAATATTCGTGTGCAAGGACCGACAACA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGAATCGGTACAGGTAGCAATGCGAGTGCCTATGTTGTTGCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGCTCTTAACCTGAAGCAGTAAAGCAATCAGAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCTCCTCAGGTTATGACCATCTGTGCAGTTCGTAATGCTGTCATCTTCGAGTCT 2238
DB 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTCTCGAATCGGTAGAGAAATTCCTGGAATGGGATTCAGGAGTGCACAGACGA 2298
DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

RESULT 3

```
US-09-697-186B-11
; Sequence 11, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-697-186B-11
```

```
Alignment Scores:
Pred. No.: 6.5e-138 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 2 Gaps: 0
```

US-10-681-086-1 (1-3156) x US-09-697-186B-11 (1-253)

```
QY 1579 GTGAATGGACCAATAATAATGACTAGACAGAAAGATGAAGATTGTTCAATGAATTAAG 1638
DB 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGATATTTGGAATAATATGCGGATCATGTTAAGCTATTGGTGTATTGGCTCTCTT 1698
DB 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGCTCAGACTGATGGCCCTATTTCGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
```

```
DB 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAAGCAGAGTTTCAGCATGATGCAACACCGGTGAGTGGAGGTGGAAGTCAATTTTGAT 1818
DB 61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
DB 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGATCATCTTAGAAGAAGTGTAT 1938
DB 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAAACTCTAAATCGGTAGAGCCCAACCGTTCACGATGCGATTTGTGCCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTTCAATATGCAAGTAAATGGCGTAATATTCGTGTGCAAGGACCGACAACA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGAATCGGTAGCAATGCGAGTGCCTATGTTGATTTGTTGCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGAGCGCTTCGCTCTTAACCTGAAGCAGTAAAGCAATCAGAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCTCCTCAGGTTATGACCATCTGTGCAGTTCGTAATGCTGTCATCTTCGACTCT 2238
DB 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTCTCTGAACTCGCTTAGAGAAATTCCTGGAATGGGATTCAGGAGTGCACAGACGA 2298
DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

RESULT 4

```
US-09-697-186B-1
; Sequence 1, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type Knt gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-09-697-186B-1
```


Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTCTCGAATCCCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpAlaGluArg 240
Qy 2299 CACGGATATATAGTGGATGTCTAAAACGCATACATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 6
US-09-697-186B-14
; Sequence 14, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-09-697-186B-14

Alignment Scores:
Pred. No.: 1.01e-131 Length: 253
Score: 1277.00 Matches: 241
Percent Similarity: 97.6% Conservative: 6
Best Local Similarity: 95.3% Mismatches: 6
Query Match: 23.1% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-14 (1-253)

Qy 1579 GTGAATGGACCAATAATACTAGACAGAGAAATGAAGATTGTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGAATATGGATAAATATGGGATCATGTTAAGGCTATTTGGTTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGCCCTATTCGGATATTGAGATGATGTGTGCATGTCACACAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
Qy 1759 GAACACAGATTCACCATGAATGACACCGGTGAGTGGAGGTGAAGTGAATTTGAT 1818
Db 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
Qy 1819 AGCGAAGAGATTCCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100
Qy 1879 GGTCAATTTTCTTATTTGGCCGATTATGATTCAGGTGATCTTACAGAAAGTGTAT 1938
Db 1938

Db 101 GlyArgPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAAACTCTAAATCGGTAGAGCCCAACGTTCCACCATGCGATTGTGCGCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTTGAATATGACAGCAAAATGGCGTAATATTCTGTGTGCAAGGACCGACACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGACTGTACAGGTAGCAATGCCAGTGCCTATGTTGATTTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGCATCTGTTATACGACGAGCGCTTCGTCTTAACATGAAGCAGTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
Qy 2179 CTTCTTCAGGTTATGACCATCTGTGCCAGTTCTGTAATGTCTGTCACTTCCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTCTCGAATCCCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpThrGluArg 240
Qy 2299 CACGGATATATAGTGGATGTCTAAAACGCATACATTT 2337
Db 241 HisGlyTyrIleValAsnValSerLysArgIleProPhe 253

RESULT 7
US-09-697-186B-15
; Sequence 15, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; CURRENT FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-09-697-186B-15

Alignment Scores:
Pred. No.: 9.95e-131 Length: 253
Score: 1268.00 Matches: 239
Percent Similarity: 97.2% Conservative: 7
Best Local Similarity: 94.5% Mismatches: 7
Query Match: 23.0% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-15 (1-253)

Qy 1579 GTGAATGGACCAATAATACTAGACAGAGAAATGAAGATTGTCATGAATTAAG 1638
Db 1938

Db	1	MetSerGlyProIleIleMetThrArgGluGluArgMetIleValHisGluIleLys	20
Qy	1639	GAACGAATATTGGATAAATATGGCGATGATGTTAAAGCTATTGGTGTTTATGGCTCTCTTT	1698
Db	21	GluArgIleLeuAspLysThrGlyAspAspValLysIleIleGlyValThrGlySerLeu	40
Qy	1699	GGTCGTGAGACTGATGGCCCTATTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG	1758
Db	41	GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu	60
Qy	1759	GAACGACAGTTCAGCCATGAATGCAGCAACCGGTGAGTGGAAAGTGGAAAGTGAATTTGAT	1818
Db	61	GlyAlaGluPheSerTyrGluIrrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr	80
Qy	1819	AGCGAAGAGATTCTACTAGATTATGTCATCTCAGGTGGAATCAGATTGGCGCCTTACACAT	1878
Db	81	SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis	100
Qy	1879	GGTCAATTTTCTCTATTTTCCGAGTTATGATTCAGTGGATACTTATAGAGAAGTGTAT	1938
Db	101	GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr	120
Qy	1939	CAAACTCTAAATCGGTAGAGGCCAAACGTTCCACGATCGGATTTGTGGCCTTATCGTA	1998
Db	121	GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal	140
Qy	1999	GAAGAGCTGTTTGAATATGCAGGCCAATGCGCTAATATTCTGTGTCGAAGGACCGACAACA	2058
Db	141	GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr	160
Qy	2059	TTTCTACCATCTTGACTGTACAGTAGCAATGGCAGGTGCCATGTTGATTTGCTGTGCAT	2118
Db	161	PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis	180
Qy	2119	CATCGCATCTGTTATACAGAGCGCTTCGGTCTTAACTGAACAGATTAGCAATCAGAT	2178
Db	181	HisArgIleCysTyrThrThrSerAlaThrValLeuThrGluAlaValLysLeuSerAsp	200
Qy	2179	CTTCTCTCAGGTATGACCATCTGTGCCAGTTCGTATGTCGGTCAACTTCCGACTCT	2238
Db	201	LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer	220
Qy	2239	GAGAAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA	2298
Db	221	GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpThrGluArg	240
Qy	2299	CACGGATATATAGTGGATGTGTCAAAACCGATACCATTT	2337
Db	241	HisGlyTrpIleValAspValSerLysArgIleProPhe	253

RESULT 8

US-09-697-186B-2
; Sequence 2, Application US/09697186B
; Patent No. 6723543

FACEBOOK NO. 0723343
: GENERAL INFORMATION:

APPLICANT: YOKOVAMA
GENERAL INFORMATION: SIEGEMIKT

APPLICANT: YUKOYAMA, SIG
APPLICANT: HOSEKI JUN

APPLICANT: HOSOKI, JUN
APPLICANT: YANO, TAKATO

APPLICANT: YANU, TAKAIO
APPLICANT: KOYAMA YOSHINOBU

APPLICANT: KIRAMITSU SEIKI
; APPLICANT: ROYAMA, YOSHINORI

APPLICANT: KURAMITSU, SEIKI
APPLICANT: KACAMIKYAMA HIROAKI

APPLICANT: KAGAMIYAMA, HIROYUKI

TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD OF TREATING OF CRYPTOCOCAL MENINGITIS IN GUinea PIG

; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA FOR THE PRODUCTION OF BIOHYDROGEN.

FILE REFERENCE: 04853.0048-000000

; CURRENT APPLICATION NUMBER: US/09/697,186B

;
; CURRENT FILING DATE: 2000-10-27

; PRIOR APPLICATION NUMBER: J

**;
PRIOR FILING**

; NUMBER OF SE

; SOFTWARE: I

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KXT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-09-697-186B-2

```

Alignment Scores:			
Pred. No.:	1.28e-130	Length:	253
Score:	1367.00	Matches:	240
Percent Similarity:	96.4%	Conservative:	4
Best Local Similarity:	94.9%	Mismatches:	9
Query Match:	23.0%	Indels:	0
DB:	2	Gaps:	0

US-10-681-086-1 (1-3156) x US-09-697-186B-2 (1-253)

1579	Qy	GTCAATGGACCAATAATAATAGCTAGAGAAGAAGAATGAAGATTCTTCATGAATAATTAAG	1638
:	Db	:::	
1	Db	MetLysGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys	20
1639	Qy	GAACGAATATTGATAAATAATGCGGATGATGTAAAGCTATTGGCTTTATGGCTCTCTT	1698
21	Db	GluArgIleLeuAspLysIstyrGlyAspAspValLysAlaIleGlyValIstyrGlySerLeu	40
1699	Qy	GGTCGTGAGACTCATGGCGCCATTCCGGATATTGACATGATGTGTCTCATGTCAACACAGAG	1758
41	Db	GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu	60
1759	Qy	GAACGAGATTCCAGCCATGAATGGACAACCGGTGAGTGAAGGTGGAAGTGAATTTTGAT	1818
61	Db	GlyAlaGluPheSerTyrGluIstThrThrGlyGluIstPheAlaGluValAsnPheTyr	80
1819	Qy	AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCGCTTACACAT	1878
81	Db	SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspIstPheLeuThrHis	100
1879	Qy	GGTCAATTTTCTCTATTTTCCGATTATTGATTCAGGTGGATACCTAGAGAAAGTGTAT	1938
101	Db	GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr	120
1939	Qy	CAAACTGCTAAATCGGTAGAAGCCCAAAAGTCCACGATGCGAATTTGTGCGCTTATCGTA	1998
121	Db	GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal	140
1999	Qy	GAAGACTGTTTGAATATCGACGCAATCGCTTAATTTCGTGTGCAAGGACGCAACA	2058
141	Db	GluGluLeuPheGluTyrAlaGlyIstPheArgAsnIleArgValGlnGlyProThrThr	160
2059	Qy	TTTCTACCATCTCTGATGTACAGGTAGCAATGGCAGGTGCCATGTTTGATGGCTGCGAT	2118
161	Db	PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis	180
2119	Qy	CATCGCATCTGTTATACGACGACGCGTCTCGCTCTTAACTGAACGAGTTAAGCAATCAGAT	2178
181	Db	HisArgIleCysIstyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp	200
2179	Qy	CTTCTCTCAGGTTATGACCATCTGTGCCAGTTCGTAATCTCTGGTCAACTTTCCGACTCT	2238
201	Db	LeuProSerGlyTyrAspHisLeuCysGlnLeuValMetSerGlyGlnLeuSerAspSer	220
2239	Qy	GAGAAACTCTGGAATCGGTAGAGAATTTCTGGAATGGGATTCCAGGAGTGGACAGAACGA	2298
221	Db	GluLysLeuLeuGluSerLeuGluAsnPheThrAsnGlyIleGlnGluIstPheThrGluArg	240
2299	Qy	CACGGATATATAGTGGATGTGTCAAAACGATACCATTT	2337
241	Db	HisGlyTyrIleValAspValSerLysArgIleProPhe	253

RESIT.T 9

RESULTS
IIS-09-697-186B-13

US-09-697-186B-13
: Sequence 13. Application US/09697186B


```
QY 1819 AGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGGAATCAGATTGGCCGCTTACACAT 1878
DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGGATACCTTAGAGAAAGATGTAT 1938
DB 101 GlyIysPhePheSerIleLeuProIleTyrAspProGlyGlyTyrLeuGluIysValTyr 120
QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGGGATTTGGCCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGGATATGCGAGCAAAATGGCGTAATATTCGTGCAAGGACCGACACACA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGGACTGACAGGTAGCAATGGCAGGTGCATGTTGATTTGGTCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACAGAGCGCTTCGGTCTTAATCGAAGCGATTTAAGCAATCAGAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
QY 2179 CTTCCTTCAGGTATGACCATCTGCGCAGTTCGTAATGTCGTGCACTTTCCGACTCT 2238
DB 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTTCTCGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGACGA 2298
DB 221 GluIysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

RESULT 11

```
US-09-697-186B-17
; Sequence 17, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KAYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; CURRENT FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-09-697-186B-17
```

```
Alignment Scores:
Pred. No.: 4,586-130 Length: 253
Score: 1262.00 Matches: 240
Percent Similarity: 96.4% Conservative: 4
Best Local Similarity: 94.9% Mismatches: 9
```

```
Query Match: 22.9% Indel: 0
DB: 2 Gaps: 0
US-10-681-086-1 (1-3156) x US-09-697-186B-17 (1-253)
QY 1579 GTGAATGACCAATAATAATGACTAGAGAAAGAAATGAAGATTTGTTTCATGAATAAAG 1638
DB 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTTGATAAATATATGGGATGATGTTAAGGCTATTTGGTGTATTTGGCTCTCT 1698
DB 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGCTCAGACTGATGGCCCTATTTCCGATATTCAGATGATGTGTGTCATGTCAACACAG 1758
DB 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGACAGATTCCACCATGAATGACCAACCGGTGAGTGGAAAGTGGAAATTTTGTAT 1818
DB 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGGAATCAGATTGGCCGCTTACACAT 1878
DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGGATCTTAGAGAAAGTGTAT 1938
DB 101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluIysValTyr 120
QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGGGATTTGGCCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGGATATGCGAGCAAAATGGCGTAATATTCGTGCAAGGACCGACACACA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGGACTGACAGGTAGCAATGGCAGGTGCATGTTGATTTGGTCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACAGAGCGCTTCGGTCTTAATCGAAGCGATTTAAGCAATCAGAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
QY 2179 CTTCCTTCAGGTATGACCATCTGCGCAGTTCGTAATGTCGTGCACTTTCCGACTCT 2238
DB 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTTCTCGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGACGA 2298
DB 221 GluIysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
RESULT 12
US-09-697-186B-18
; Sequence 18, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KAYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; CURRENT FILING DATE: 2000-10-27
```

;; PRIOR APPLICATION NUMBER: JP 309616/1999
;; PRIOR FILING DATE: 1999-10-29
;; NUMBER OF SEQ ID NOS: 20
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 18
;; LENGTH: 253
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
;; OTHER INFORMATION: obtained by introduction of point mutation into
;; OTHER INFORMATION: wild type KMT gene of Staphylococcus aureus and
;; OTHER INFORMATION: its expression

US-09-697-186B-18

Alignment Scores:
Pred. No.: 4,59e-130 Length: 253
Score: 1262.00 Matches: 240
Percent Similarity: 96.4% Conservatives: 4
Best Local Similarity: 94.9% Mismatches: 9
Query Match: 22.9% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-18 (1-253)

```
Qy 1579 GTGAATGGACCAATAATAAGTACAGAGAAAGAAATGAAGATTGTTTCATGAATAAAG 1638
Db 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20

Qy 1639 GAACGAATATGGATAATATGGGATGATGTTAAGGCTATTGTTGTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40

Qy 1699 GGTGCTCAGACTGATGGCCCTATTTCGATATTTCAGATGATGTCGTGCATGCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60

Qy 1759 GAAGCAGAGTTTCAGCCCAATGAATGGAACACCGGTGAGTGGAGGTGGAAGTGAATTTGAT 1818
Db 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80

Qy 1819 AGCGAAGAGATTCTACTAGATTATGATTCAGTGCAGTGAATCAGATGGCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100

Qy 1879 GGTCAATTTTCTCTATTTGCGCATTTATGATTCAGTGCAGTGAATTTGATGAAAGTGTAT 1938
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspThrGlyGlyTyrLeuGluLysValTyr 120

Qy 1939 CAAACTGCTAAATCGGTAGAACCCCAACCGTTCACGATGCGATTTCGTCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140

Qy 1999 GAAGAGCTGTTGAATATGAGCAAAATGCGTAAATATTCGTGTGCAAGGACCGACAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlnProThrThr 160

Qy 2059 TTTCTACCATCTTACTGTACAGTAGCAATGCGAGTGCCTGCTGTTGATTGCTGCTCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

Qy 2119 CATGCGATCTGTTATACGACGAGCGCTTCGGTCTTAATCTGAAGCAGTGAATGAACATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200

Qy 2179 CTTCCTTCAGTTATGACCATCTGTGCGAGTTGCTGAATGCTGTGCTCACTTCGACTCT 2238
Db 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

Qy 2239 GAGAAATCTTCGGAATCGTAGCAATTTCTGGAATCGGATTCAGGATTCAGGAGTGCACAGACA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGlnTrpThrGluArg 240
```

```
Qy 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

RESULT 13

US-09-697-186B-16

;; Sequence 16, Application US/09697186B
;; Patent No. 6723543
;; GENERAL INFORMATION:
;; APPLICANT: YOKOYAMA, SIGEYUKI
;; APPLICANT: HOSEKI, JUN
;; APPLICANT: YANO, TAKATO
;; APPLICANT: KOYAMA, YOSHINORI
;; APPLICANT: KURAMITSU, SEIKI
;; APPLICANT: KAGAMIYAMA, HIROYUKI
;; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
;; FILE REFERENCE: 04853.0048-00000
;; CURRENT APPLICATION NUMBER: US/09/697,186B
;; CURRENT FILING DATE: 2000-10-27
;; PRIOR APPLICATION NUMBER: JP 309616/1999
;; PRIOR FILING DATE: 1999-10-29
;; NUMBER OF SEQ ID NOS: 20
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 16
;; LENGTH: 253
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
;; OTHER INFORMATION: obtained by introduction of point mutation into
;; OTHER INFORMATION: wild type KMT gene of Staphylococcus aureus and
;; OTHER INFORMATION: its expression

US-09-697-186B-16

Alignment Scores:
Pred. No.: 7,62e-130 Length: 253
Score: 1260.00 Matches: 237
Percent Similarity: 96.0% Conservatives: 6
Best Local Similarity: 93.7% Mismatches: 10
Query Match: 22.8% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-16 (1-253)

```
Qy 1579 GTCAATGGACCAATAATAAGTACAGAGAAAGAAATGAAGATTGTTTCATGAATAAAG 1638
Db 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20

Qy 1639 GAACGAATATGGATAATATGGGATGATGTTAAGGCTATTGTTGTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40

Qy 1699 GGTGCTCAGACTGATGGCCCTATTTCGATATTTCAGATGATGTCGTGCATGCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60

Qy 1759 GAAGCAGAGTTTCAGCCCAATGAATGGAACACCGGTGAGTGGAGGTGGAAGTGAATTTGAT 1818
Db 61 GluThrGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80

Qy 1819 AGCGAAGAGATTCTACTAGATTATGATTCAGTGCAGTGAATCAGATTCGCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100

Qy 1879 GGTCAATTTTCTCTATTTGCGCATTTATGATTCAGTGCAGTGAATTTGATGAAAGTGTAT 1938
Db 101 GlyLysPhePheSerIleLeuProIleTyrAspThrGlyGlyTyrLeuGluLysValTyr 120

Qy 1939 CAAACTGCTAAATCGGTAGAACCCCAACCGTTCACGATGCGATTTCGTCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
```

QY 1999 GAAGAGCTGTTGTAATATGACAGCAAAATGGGCTAATATTCGTGTGCAAGGACCCGACAACA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleAaGValGlnGlyProLeuThr 160
QY 2059 TTTCTACCATCTTGACTGTACAGGTAGCAATAGGAGTGCATGTTGATGGTCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTAGAGCGATTAAGCAATCAGAT 2178
DB 181 HisArgIleCysTyrThrGlyAlaSerValLeuThrGluAlaValArgGlnProAsp 200
QY 2179 CTTCCTCAGGTTATGACCATCTGTGCGAGTTTCGTAATGTCTGGTCAACTTTCGAGCTCT 2238
DB 201 LeuProGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTTCTGGAATCCTGAGAAATTCGGAATGGGATTCAGAGTGCAGAGAACA 2298
DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAenGlyIleGlnGluTrpAlaGluArg 240
QY 2299 CACGGATATAGTGATGTGTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 14

US-09-697-186B-19
; Sequence 19, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: Obtained by introduction of point mutation into
; OTHER INFORMATION: wild type Knt gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression

US-09-697-186B-19

Alignment Scores:
Pred. No.: 2,11e-129 Length: 253
Score: 1256.00 Matches: 239
Percent Similarity: 96.0% Conservative: 4
Best Local Similarity: 94.5% Mismatches: 10
Query Match: 22.8% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-19 (1-253)

QY 1579 GTGAATGGACCAATAATAGTACTAGAGAAGAAGAAATGAAGATGTTTCATGAATTAAG 1638
DB 1 MetaAenGlyProIleLeuMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATGGATAAATATGGGATGATGTTAAGGCTATGTTGTTTATGGCTCTCTT 1698
DB 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTGCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
DB 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACAGAGTTTACCCATGAATGCAACACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGAAT 1818
DB 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
QY 1819 AGCCAAGAGATCTTACTAGATTATGATCTCAGAGTGGAAATCAGATTCGCGCCCTTACACAT 1878
DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTCGCCCATTTATGATTTCAGTGTGATATCTTAGAGAAAGTGTAT 1938
DB 101 GlyLysPhePheSerIleLeuProIleTyrAspProGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAAACTGCTAAATCGGTAGAACCCAAACGTTTCACGATGCGATTTGTGCGCCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGTAATATGACGAGCAAAATGGGCTAATATTCGTGTGCAAGGACCCGACAACA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleAaGValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGACTGTACAGGTAGCAATGGCAGTGCCATGTTGATTCGTCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAACAATCAGAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
QY 2179 CTTCCTCAGGTTATGACCATCTGTGCGAGTTTCGTAATGTCTGTCACACTTTCGACTCT 2238
DB 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspPro 220
QY 2239 GAGAACTTCTGGAATCGTAGAATTTCTGGAATGGGATTCAGAGTGCAGAGAACA 2298
DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAenGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATAGTGATGTGTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 15

US-09-697-186B-3
; Sequence 3, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: Obtained by introduction of point mutation into
; OTHER INFORMATION: wild type Knt gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression

US-09-697-186B-3

Alignment Scores:
Pred. No.: 1.21e-125 Length: 253
Score: 1222.00 Matches: 231
Percent Similarity: 93.7% Conservative: 6
Best Local Similarity: 91.3% Mismatches: 16
Query Match: 22.1% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-3 (1-253)

Qy	1579	GTGAATCGACCAATATAATGACTAGAGAGAAAGATGATGTTCAATGAATTAAG	1638
Db	1	MetLysGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys	20
Qy	1639	GAACGAATATGGTAATATGGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTT	1698
Db	21	GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu	40
Qy	1699	GGTCGTCAGACTGATGGCCCTATTTCGATATTGAGATGATGTGTGTCATGTCAACAGAG	1758
Db	41	GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValLeuSerThrGlu	60
Qy	1759	GAACGAGAGTTTCAGCCATGAATGACCAACCGGTGAGTGGAGGTGGAAGTCAATTTTGTAT	1818
Db	61	GlyValGluPheSerTyrGluTyrThrThrGlyGluTyrLysAlaGluValAsnPheTyr	80
Qy	1819	AGCGAAGAGATTCTACTAGATTATGCACTCTCAGGTGGAATCAGATTCGCCGCTTACACAT	1878
Db	81	SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTyrProLeuThrHis	100
Qy	1879	GGTCAATTTTCTCTATTTCGCGATTATGATTTCAGTGGTACTTTAGAGAAAGTGTAT	1938
Db	101	GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr	120
Qy	1939	CAAACTGCTAAATCGGTAGAGCCCAACCGTTCACGATGCGATTTGTGCCCTTTATCGTA	1998
Db	121	GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal	140
Qy	1999	GAAGAGCTGTTTGAATATGCGGCAATATGGCGTAATATTCGTGTGCAAGGACCGACAACA	2058
Db	141	GluGluLeuPheGluTyrAlaGlyLysTyrPargAsnIleArgValGlnGlyProThrThr	160
Qy	2059	TTTCTACCATCTTGACTGTACAGTAGCAGTGCAGTGCATGTTGATTGGTCTGCAT	2118
Db	161	PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis	180
Qy	2119	CATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGAT	2178
Db	181	HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp	200
Qy	2179	CTTCCTTCAGGTTATGACCATCTGTGCGAGTTCGTATGCTGTGCTCAACTTCCGACTCT	2238
Db	201	LeuProProGlyTyrValGlnLeuCysGlnLeuValMetSerGlyGlnLeuSerAspPro	220
Qy	2239	GAGAACTTCTGGAATCGCTAGAGATTTCTGGAATTCGAGATTCAGGAGTGGACAGACGA	2298
Db	221	GluLysLeuLeuGluSerLeuGluAsnPheTyrAsnGlyValGlnGluTyrPalGluArg	240
Qy	2299	CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT	2337
Db	241	HisGlyTyrIleValAspValSerLysArgIleProPhe	253

Search completed: February 10, 2006, 03:03:33
Job time : 111 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2006, 03:49:13 ; Search time 87.9 Seconds
(without alignments)
3000.385 Million cell updates/sec

Title: us-10-681-086-1

Perfect score: 5517

Sequence: 1 ggaatccacgaggttaacagc.....atcggtatgatgcgaattc 3156

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/abs/ABSSWEB_spool/US10681086/runat_09022006_140527_11976/app_query.fasta_1
-DB=Published Applications_AA_Main -QPMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pto -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10681086 @CGN 1.1 307 @runat_09022006_140527_11976 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -RSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_Main:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1836	33.3	389	4	US-10-369-493-23255
2	1333	24.2	253	4	US-10-320-800-75
3	1333	24.2	253	4	US-10-718-628-11
4	1318	23.9	253	4	US-10-718-628-1
5	1282	23.1	253	4	US-10-718-628-12
6	1277	23.2	253	4	US-10-718-628-14
7	1268	23.0	253	4	US-10-718-628-15
8	1267	23.0	253	4	US-10-718-628-2
9	1267	23.0	253	4	US-10-718-628-13
10	1265	22.9	253	4	US-10-718-628-20
11	1262	22.9	253	4	US-10-718-628-17

12	1262	22.9	253	4	US-10-718-628-18
13	1260	22.8	253	4	US-10-718-628-16
14	1256	22.8	253	4	US-10-718-628-19
15	1222	22.1	253	4	US-10-718-628-3
16	942	17.1	387	4	US-10-033-078-6
17	942	17.1	387	4	US-10-763-333-6
18	851	15.4	338	4	US-10-033-078-8
19	851	15.4	338	4	US-10-763-933-8
20	843	15.3	392	4	US-10-369-493-23135
21	828	15.0	321	4	US-10-282-122A-71209
22	828	15.0	327	4	US-10-724-972A-4421
23	809	14.7	332	4	US-10-282-122A-45833
24	793	14.4	338	3	US-09-815-242-5784
25	793	14.4	335	3	US-09-815-242-12780
26	793	14.4	335	3	US-09-815-242-13125
27	793	14.4	335	4	US-10-282-122A-44022
28	778	14.1	321	4	US-10-282-122A-71676
29	767	13.9	371	4	US-10-369-493-22995
30	712	12.9	392	4	US-10-369-493-13494
31	712	12.9	395	4	US-10-369-493-17471
32	711	12.9	396	4	US-10-282-122A-46084
33	710	12.9	388	4	US-10-369-493-16649
34	705	12.8	395	3	US-09-815-242-12441
35	704	12.8	385	3	US-09-815-242-5575
36	704	12.8	396	4	US-10-282-122A-58006
37	700	12.7	393	4	US-10-369-493-18251
38	699	12.7	324	4	US-10-282-122A-45547
39	699	12.7	398	4	US-10-282-122A-71674
40	699	12.7	400	4	US-10-369-493-20208
41	698.5	12.7	372	4	US-10-369-493-1058
42	695.5	12.6	400	5	US-10-501-282-2370
43	692	12.5	398	4	US-10-369-493-21684
44	689	12.5	391	4	US-10-369-493-16710
45	688	12.5	402	4	US-10-369-493-18774

ALIGNMENTS

RESULT 1

US-10-369-493-23255
; Sequence 23255, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCES: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23255
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Bacillus subtilis
; US-10-369-493-23255

Alignment Scores:
Pred. No.: 3 6e-173 Length: 389
Score: 1836.00 Matches: 359
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 33.3% Indels: 0
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-369-493-23255 (1-389)

QY 22 TTGAAGATTGCTCTGGTTAAACGAGCGGTAGACAGATGAAGAAGCCGGGTACAT 81

```
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1 MetLysIleAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20
QY      82 CGTAACCTGCGGTCAATGATGGAGCGCGGTTCAGAGAGAAATATTGATGGCGAAAT 141
Db      21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
QY      142 CAACGGCTCGTCTCTCAACAATATTATTAGGGCTCGCAAGCGATAGACGTTTTCATCGAT 201
Db      41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgArgLeuIleAsp 60
QY      202 GCAGCCCAACAGCATTCAGCAATTTTCGGGACAGGAACACGCGTTCACGTTTAAACGACA 261
Db      61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
QY      262 GGCAATTCGGTCTGCGCATGAAGAAGCTAGAAAGAGATTGCCAGCTTTAAACTGCACAAA 321
Db      81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100
QY      322 GCGGCCCTGCTGTTTTCAGCGGTTTACTTGGCCAATGTCGGTGTCTCTTTCATCCTTGCCA 381
Db      101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
QY      382 GAAAGGAGAGATGTCATTTTAAGTGACCAGCTCAATCATGCAAGTATGATCGCGCTGC 441
Db      121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
QY      442 CGACTTCTAGGCTGATACAGTGTGTTTATCGGCATATTGATATGAATGATCTTGAAGAAC 501
Db      141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160
QY      502 AAGCTGAATGAAACACACAGCGTTATCAGCGCGCTTTTATCGTAACAGCGAGTATTACGC 561
Db      161 LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
QY      562 ATGATGACCAATCGCCCTCTTGATCAGATCATCTCACTTGGCAACCGCTATCATGCC 621
Db      181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTyrHisAla 200
QY      622 TTCGTGTCGTGATGATGCCACCAAGAGAGTTTGGCGGATTCGGGACAGGAACG 681
Db      201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
QY      682 AGTGAATACTTTGGTGTGTTTCCGACATGTTATCGGCACCTTAAGCAAAAGCTGTTCGC 741
Db      221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
QY      742 CGGGAAGGAGGTTTTCGGCAGGATCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db      241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
QY      802 AGAACATTTATCTTCAACCGCTATTTCGGCCAGCCAGCTGTGGCGTGTCTCAGAGGCT 861
Db      261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaAlaHisGluAla 280
QY      862 TTCACATCATTTGAAGCCAGGAGGAAAAACGACAGCTTTTATTTCTATATCAGCATG 921
Db      281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
QY      922 ATCAGAACCAAGCTGGAAGAATATGGTTTATGTGTTGAAAGGAGATCACACACCGATTAT 981
Db      301 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 320
QY      982 CCTGTAGTCATTTGGCGATGCCATAAACCGTCTCTATTGCTGAAAAAATCTCAGGGCAAG 1041
Db      321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
QY      1042 GGAATTTTATGCTCCGTGCCATTTCGGCCGCAACCGTTGCGCGGTGGAAGCCGATTCCGA 1101
Db      341 GlyIleTyrAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 360
```

RESULT 2

US-10-320-800-75

```
; Sequence 75, Application US/10320800
; Publication No. US20030215469A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, ANDREW
; APPLICANT: GORRINGE, ANDREW
; APPLICANT: HUDSON, MICHAEL
; APPLICANT: REDDIN, KAREN
; TITLE OF INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE
; FILE REFERENCE: 1581.0790001
; CURRENT APPLICATION NUMBER: US/10/320,800
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: PCT/GB99/03626
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 253
; TYPE: PRT
; ORGANISM: unidentified
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: kanamycin resistance
; US-320-800-75

Alignment Scores:
Pred. No.:      3,78e-123      Length:      253
Score:          1333.00      Matches:      252
Percent Similarity: 100.0%      Conservative: 1
Best Local Similarity: 99.6%      Mismatches: 0
Query Match:      24.2%      Indels:      0
DB:              4          Gaps:      0

US-10-681-086-1 (1-3156) x US-10-320-800-75 (1-253)
```

```
QY      1579 GTGAATCGACCAATAATAATGACTAGAGAGAAAGATGAAGATTGTTTCATGAATTAAG 1638
Db      1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY      1639 GAACGAATATTGGATAAATATGGGATGATGGGATGATGTTAAAGCTATTGCTGTTTATGGCTCTCTT 1698
Db      21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
QY      1699 GGTGCTCAGACTGATGGGCCCTATTTCGGATATTTCGAGATGATGTGTCATGTCACACAGAG 1758
Db      41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY      1759 GAACGAGATTTCACCCATGAATGGACAAACCGTGAGTGGAAAGTGGAAAGTGAATTTTCAT 1818
Db      61 GluAlaGluPheSerHisGluTrpThrThrGlyLysValGluValAsnPheAsp 80
QY      1819 AGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878
Db      81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY      1879 GGTCAATTTTCTCTATTTTGCGGATTTATGATTTCAGTGGATACTTAGAGAAAGTGTAT 1938
Db      101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY      1939 CAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGGGATTTGTGCCCTTATCGTA 1998
Db      121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY      1999 GAACGAGCTGTTTGAATATGACGGCAAAATGGCGTAATATTTCGTGTGCAAGGACCGACAACA 2058
Db      141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY      2059 TTCTTACCATCTCTTGACTGTACAGGTAGCAATGGCAGTGCATGTTGATTGGTCTGCAT 2118
Db      161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY      2119 CATCGATCTGTTATACAGACGAGCGCTTCGGTCTTAAGTAAAGCAGTAAAGCAATCAGAT 2178
Db      181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
```

```
QY 2179 CTTCTTCAGGTTATGACCAATCTGTGCCAGTTCGTAATGTGTGTCACAACTTTCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAAATCTTCGGAATCGCTAGAGAAATCTTCGGAATCGGATTCAGGATGACAGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerIysArgIleProPhe 253

RESULT 3
US-10-718-628-11
; Sequence 11, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; CURRENT APPLICATION NUMBER: US/10718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-718-628-11

Alignment Scores:
Pred. No.: 3.78e-123 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.3% Indels: 0
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-11 (1-253)
QY 1579 GTGAATGGACCAATAATAATGACCTAGACAGAAAGAAATGAGATTGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTGGATAATATGCGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTCTGACAGCTGATGGCCCTATTTCGGATATTCAGATGATGTGTGTCATGTCACACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGAGATTACCCATGAATGACACACCGGTGAGGTGGAGTGAATTTGAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGATTCAGGTGGAAATCAGATTGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspIysAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTTGCCGATTATTGATTCAGGTGGATCTTAGAGAAAGTGTAT 1938
```

```
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAAATCGCTAAATCGGTAGAACCCCAAAACGCTTCACGATCGGATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAAGCTGTTTGAATATGACGCAAAATGCGGTAATAATTCTGTGTGCAAGACCGACACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGACTGTACAGGTAGCAATGCGAGTGCATGTTGATTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGAGCGTTCGCTTAACTGAAGCAGTAAACCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCTTCAGGTTATGACCAATCTGTGCCAGTTCGTAATGTGTGTCACAACTTTCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAAATCTTCGGAATCGCTAGAGAAATTTCTGGAATCGGATTCAGGATGACAGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 4
US-10-718-628-1
; Sequence 1, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; CURRENT APPLICATION NUMBER: US/10718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-1

Alignment Scores:
Pred. No.: 1.18e-121 Length: 253
Score: 1318.00 Matches: 250
Percent Similarity: 99.2% Conservative: 1
Best Local Similarity: 98.8% Mismatches: 2
Query Match: 23.9% Indels: 0
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-1 (1-253)
```

```
Qy 1579 GTGAATGACCAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAG 1638
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGAATATTGGATAAATATGCGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTT 1698
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGCCCTATTTCGATATTCGATGATGATGATGATGATGATGATGAT 1758
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60
Qy 1759 GAAGCAGAGTTCAGCCATTAATGACCAACCGGTGATGCGAAGTGAAGTGAATTTTCAT 1818
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheTyr 80
Qy 1819 AGCGAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTTCGCTATTTTCGATTTTCAGGTGATCTTAGAGAAAGTGTAT 1938
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCGATTGTGCGCTTATCGTA 1998
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTGTAATATGCGAGGCAAAATGGCGTAATATTCGTGTCGAAGGCCGACA 2058
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGACTGACGTACGATGCGATGCGATGCGATGCGATGCGATGCGATGCG 2118
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGATCTGTATACGACGAGCCGCTTCGGTCTTAATGAGAGAGTGAAGCAATACAT 2178
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
Qy 2179 CTTCCTTCAGGTATGACCATCTGTCGCAATTCGGAATGCGATGCGATGCGATGCGATGCG 2238
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTCTCGAATCGCTAGACATCTGTCGCAATTCGGAATGCGATGCGATGCGATGCG 2298
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy 2299 CACGATATATAGTGGATGTGTCAAAACCGCATACCATTT 2337
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

RESULT 5

```
US-10-718-628-12
; Sequence 12, Application US/10718628
; Publication No. US2004007016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
```

```
/ SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-12
```

```
Alignment Scores:
Pred. No.: 4.56e-118 Length: 253
Score: 1282.00 Matches: 241
Percent Similarity: 97.6% Conservative: 6
Best Local Similarity: 95.3% Mismatches: 6
Query Match: 23.2% Indels: 0
DB: 4 Gaps: 0
```

US-10-681-086-1 (1-3156) x US-10-718-628-12 (1-253)

```
Qy 1579 GTGAATGACCAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValTyrGluIleLys 20
Qy 1639 GAACGAATATTGGATAAATATGCGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGCCCTATTTCGATATTCGATGATGATGATGATGATGATGATGAT 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60
Qy 1759 GAACGAGAGTTCAGCCATTAATGACCAACCGGTGATGCGAAGTGAAGTGAATTTTCAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
Qy 1819 AGCGAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTTCGCGATTTATGATTTCAGGTGATCTTAGAGAAAGTGTAT 1938
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120
Qy 1939 CAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCGATTGTGCGCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTGTAATATGCGAGGCAAAATGGCGTAATATTCGTGTCGAAGGCCGACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGACTGACGTAGCAATGCGATGCGATGCGATGCGATGCGATGCGATGCG 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGATCTGTATACGACGAGCCGCTTCGGTCTTAATGAGAGAGTGAAGCAATACAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaLeuLysGlnSerAsp 200
Qy 2179 CTTCCTTCAGGTATGACCATCTGTCGCAATTCGGAATGCGATGCGATGCGATGCGATGCG 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTCTCGAATCGCTAGACATTTCTCGAATGCGATGCGATGCGATGCGATGCGAGAAC 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpAlaGluArg 240
Qy 2299 CACGATATATAGTGGATGTGTCAAAACCGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

```
RESULT 6
US-10-718-628-14
; Sequence 14, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-14
Alignment Scores:
Pred. No.: 1,44e-117 Length: 253
Score: 1277.00 Matches: 241
Percent Similarity: 97.6% Conservative: 6
Best Local Similarity: 95.3% Mismatches: 6
Query Match: 23.1% Indels: 0
DB: 4 Gaps: 0
US-10-681-086-1 (1-3156) x US-10-718-628-14 (1-253)
QY 1579 GTCAATGCGCAATAATAAGTACTAGAGAAAGATGATGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTTCCTCTCT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GTTCGTGAGACTGATGGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGAGAGTTCAGCCATGAATGACACACCGTGCAGTGGAGGTGAAGTGAATTTGAT 1818
Db 61 GlyAlaGluPheSerTyrGluTrpThrGlyGluTrpLysAlaGluValAsnPheTyr 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTTATTTCGCGATTTATGATTTCAGTGGATCTTAGAGAAAGTGTAT 1938
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAACACTGTAATCGGTAGAGCCCAACGTTCCACCATGCGATTTCGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTTGAATATGACGGCAATGCGCGTAATATTCTGTGTGCAAGGACCGACA 2058
```

```
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGACTGTACAGGTAGCAATGCGAGGTGCCATGTTGATTGTTCTCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGCGCTTCGCTTAACTGAAGCAGTGAACATCATGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
QY 2179 CTTCTTCAGGTATGACCATCTGTGCCAGTTCGTAACTGCTGTCGTCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAAACTCTCGAATCGCTAGAGAATTCTCGAATGGGATTTCAGGATGCGACAGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAsnValSerLysArgIleProPhe 253
RESULT 7
US-10-718-628-15
; Sequence 15, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-15
Alignment Scores:
Pred. No.: 1,13e-116 Length: 253
Score: 1268.00 Matches: 239
Percent Similarity: 97.2% Conservative: 7
Best Local Similarity: 94.5% Mismatches: 7
Query Match: 23.0% Indels: 0
DB: 4 Gaps: 0
US-10-681-086-1 (1-3156) x US-10-718-628-15 (1-253)
QY 1579 GTCAATGCGCAATAATAAGTACTAGAGAAAGATGATGTTTCATGAATTAAG 1638
Db 1 MetSerGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTTCCTCTCT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
```



```

Qy 2119 CATCGCATCTGTTATACGACGACGGCTTCGCTCTTAACGACGAGTTAAGCAATCAAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaLeuValLeuThrGluAlaValIleValSerAsp 200
Qy 2179 CTTCCCTTCAGGTTATGACCAATCTGTGCAGTTCGTAAATGCTCGTCAACTTTCCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisIleCysGlnLeuValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAAACTCTCGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGGACAGAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAenPheTIPAsnGlyIleGlnGluTTPThrGluArg 240
Qy 2299 CACGGATATATAGTGGATGTGTCAAAACGATACCAATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 10
US-10-718-628-20
; Sequence 20, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCES: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10/718,628
; PRIOR FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-20

Alignment Scores:
Pred. No.: 2,25e-116 Length: 253
Score: 1265.00 Matches: 240
Percent Similarity: 96.8% Conservativeness: 5
Best Local Similarity: 94.9% Mismatches: 8
Query Match: 22.9% Indels: 0
DB: Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-20 (1-253)

Qy 1579 GTGAATGGACCAATAATAATAGCTAGAGAGAAGAATGAAGATTGTTCAATAAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGATATTCGATTAATATGGGATGATGTTAAGCGCTATTGGTGTGTTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValIleAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTGACACTCATGGCCCTATTCCGATATTGAGATGATGTGTCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValLeuSerThrGlu 60
Qy 1759 GAACGACAGATTACCGCATGAATCGAACCAACCGGTGAGTGGAGGTGGAAGTGAATTTGAT 1818

```

```
Db 61 GluAlaGluPheSerTyxGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyx 80
Qy 1819 AGCGAGAGATCTACTAGATTATGCATCTCAGGTGGAAATCAGATTGCCCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyxAlaSerArgValGluProAspTrpProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTTCGCGATTATGATTCCAGTGGATATCTAGAGAAAGTGTAT 1938
Db 101 GlyLysPhePheSerIleLeuProIleTyxAspProGlyGlyTyxLeuGluLysValTyx 120
Qy 1939 CAAACTGCTAAATCGGTAGAACCCAAACGTTCCACGATGCGATTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTTGAATATGCGAGGCAATGGCGTAATATTCGTGTCAGAGGCCGACACA 2058
Db 141 GluGluLeuPheGluTyxAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCATGTTGATTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGACTCTGTTATPACGACGAGCGCTTCGTCTTAACTGAAGCAAGTAAACCAATCAGAT 2178
Db 181 HisArgIleCysTyxThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
Qy 2179 CTTCCTTCAGGTTATGACCATCTGCGCAGTTCGTAATGTTCTGTCACACTTCCGACTCT 2238
Db 201 LeuProSerGlyTyxValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTTCCTGGAATCCGCTAGAGAAATTCCTGGAAATGGGATTCCAGGAGTGACAGAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy 2299 CACGGATATATAGTGGATGTCTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyxIleValAspValSerLysArgIleProPhe 253
```

RESULT 11

US-10-718-628-17

; Sequence 17, Application US/10718628

; Publication No. US20040077016A1

; GENERAL INFORMATION:

; APPLICANT: YOKOYAMA, SIGEYUKI

; APPLICANT: HOSEKI, JUN

; APPLICANT: YANO, TAKATO

; APPLICANT: KAYAMA, YOSHINORI

; APPLICANT: KURAMITSU, SEIKI

; APPLICANT: KAGAMIYAMA, HIROYUKI

; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD

; FILE REFERENCE: 04853.0048-00000

; CURRENT APPLICATION NUMBER: US/10/718,628

; CURRENT FILING DATE: 2003-11-24

; PRIOR APPLICATION NUMBER: US/09/697,186B

; PRIOR FILING DATE: 2000-10-27

; PRIOR APPLICATION NUMBER: JP 309616/1999

; PRIOR FILING DATE: 1999-10-29

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 17

; LENGTH: 253

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme

; OTHER INFORMATION: obtained by introduction of point mutation into

; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and

; OTHER INFORMATION: its expression

US-10-718-628-17

Alignment Scores:

Pred. No.: 4,48e-116 Length: 253

Score: 1262.00 Matches: 240

Percent Similarity: 96.4% Conservative: 4

Best Local Similarity: 94.9% Mismatches: 9

Query Match: 22.9% Indels: 0

DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-17 (1-253)

Qy 1579 GTGAATGACCAATAATAATGACTAGAGNAGAAATGAGATTGTTTCATGAAATTAAG 1638

Db 1 MetAsnGlyProIleIleMetThrArgGluArgMetLysIleValHisGluIleLys 20

Qy 1639 GAACGAATATTGGATAAATATGGCGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCT 1698

Db 21 GluArgIleLeuAspLysTyxGlyAspAspValLysAlaIleGlyValTyxGlySerLeu 40

Qy 1699 GGTGCTCAGACTGATGGGCCCTATTTCGATATATCAGATGATGTGTGCATGTTCAACAGAG 1758

Db 41 GlyArgGlnThrAspGlyProTyxSerAspIleGluMetMetCysValMetSerThrGlu 60

Qy 1759 GAACGAGTTTCACCATGATGACAAACCGTGCAGTGGAGGTGGAGTGAATTTTCAT 1818

Db 61 GlyAlaGluPheSerTyxGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyx 80

Qy 1819 AGCGAGAGATCTTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878

Db 81 SerGluGluIleLeuLeuAspTyxAlaSerArgValGluSerAspTrpProLeuThrHis 100

Qy 1879 GGTCAATTTTCTCTATTTCGCGATTATGATTTCAGTGGATCTTAGAGAAAGTGTAT 1938

Db 101 GlyArgPhePheSerIleLeuProIleTyxAspProGlyGlyTyxPheGluLysValTyx 120

Qy 1939 CAACTCTTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTGTGCCCTTATCGTA 1998

Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140

Qy 1999 GAAGAGCTGTTTGAATATGCGAGGCAATGGCGTAATATTCGTGTCAGAGGCCGACACA 2058

Db 141 GluGluLeuPheGluTyxAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

Qy 2059 TTTCTACCATCTTTCAGTGTACAGGTAGCAATGCGAGTGCATGTCGATTTGTTGTTGCTGAT 2118

Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

Qy 2119 CATCGACTCTGTTATPACGACGAGCGCTTCGTCTTAACTGAAGCAGTAAAGCAATCAGAT 2178

Db 181 HisArgIleCysTyxThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200

Qy 2179 CTTCCTTCAGGTTATGACCATCTGTCGAGTTCGTAATGTTCTGTCACACTTTCGACTCT 2238

Db 201 LeuProSerGlyTyxValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

Qy 2239 GAGAACTTCCTGGAATCGGTAGAGAAATTCCTGGAAATGGGATTCCAGGAGTGACAGAACGA 2298

Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240

Qy 2299 CACGGATATATAGTGGATGTCTCAAAACGCATACCATTT 2337

Db 241 HisGlyTyxIleValAspValSerLysArgIleProPhe 253

RESULT 12

US-10-718-628-18

; Sequence 18, Application US/10718628

; Publication No. US20040077016A1

; GENERAL INFORMATION:

; APPLICANT: YOKOYAMA, SIGEYUKI

; APPLICANT: HOSEKI, JUN

; APPLICANT: YANO, TAKATO

; APPLICANT: KAYAMA, YOSHINORI

; APPLICANT: KURAMITSU, SEIKI

; APPLICANT: KAGAMIYAMA, HIROYUKI

; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD

; FILE REFERENCE: 04853.0048-00000

; CURRENT APPLICATION NUMBER: US/10/718,628

; CURRENT FILING DATE: 2003-11-24

; PRIOR APPLICATION NUMBER: US/09/697,186B

; PRIOR FILING DATE: 2000-10-27

; PRIOR APPLICATION NUMBER: JP 309616/1999

; PRIOR FILING DATE: 1999-10-29

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 17

; LENGTH: 253

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme

; OTHER INFORMATION: obtained by introduction of point mutation into

; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and

; OTHER INFORMATION: its expression

US-10-718-628-17

Db 101 GlyLysPheSerIleLeuProIleThrGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAACTGCTAAATCGGTAGAGCCCAACGGTTCACGATGGATTTGGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTTGAATATGACGAGCAATGGCGTAAATTCGTGTGCAAGGACCCGACAAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysIlePheGlnIleArgValGlnGlyProLeuThr 160
Qy 2059 TTTCTACCATCTTGTACGTAGCATGAGTGGCGTTCCTTAAGCAAGCAATGATCGAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATGCTCTCTTATACAGACAGCGCTTCGGTCTTAAGCAAGCAATGATCGAT 2178
Db 181 HisArgIleCysTyrThrGlyAlaSerValLeuThrGluAlaValArgGlnProAsp 200
Qy 2179 CTTCCTTCAGGTATGACCATCTGCGCAGTTCGTAATGCTGTGTCACACTTTCGCACTCT 2238
Db 201 LeuProProGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTTCTCGAATCCCTAGACAAATTCGGAATGGGATTCAGGAGTGGACAGAAACA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpAlaGluArg 240
Qy 2299 CACGATATATAGTGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 14

US-10-718-628-19
; Sequence 19, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10/718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type Knt gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression

US-10-718-628-19
Alignment Scores:
Pred. No.: 1,786-115 Length: 253
Score: 1256.00 Matches: 239
Percent Similarity: 96.0% Conservative: 4
Best Local Similarity: 24.5% Mismatches: 10
Query Match: 22.8% Indels: 0
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-19 (1-253)

Qy 1579 GTGAATGACCAATAATAATGACTAGAGAAAGAAATGAAGATTGTTTCATGAATAAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGAATATTGGATAAATATGGGATCATGTTTAAGGCTATTGCTGTTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGCCCTTATTCGATATTGAGATGATGCTGTCATGTCACACAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
Qy 1759 GAACACAGATTTCACCCATGAATGACAAACCGTTCAGTGGAGGTGGAAGTCAATTTTCAT 1818
Db 61 GlyAlaGluPheSerTyrGluTrpThrGlyGluTrpLysAlaGluValAsnPheTyr 80
Qy 1819 AGCAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTyrProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTGTCGCAATTATGATTCAGTGGATACATTAGAGAAAGTGTAT 1938
Db 101 GlyLysPhePheSerIleLeuProIleTyrAspProGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAAACTCTAAATCGGTAGAGCCCAACCGTTCACGATGCGATTGTCGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTTGAATATGACGAGCAATGCGCGTAAATTCGTGTGCAAGGACCCGACAAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGTACTGATGACGATGACGATGCGCGTTCAGTGGATGTTGATTTGGTCTCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATGCTCTCTTATACGACGAGCGTTCGGTCTTAAGCAAGCAATGATGAACAAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
Qy 2179 CTTCCTTCAGGTTATGACCATCTGTCGCGTTCGTAATGCTGTCATGCTCAACTTTCGCACTCT 2238
Db 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspPro 220
Qy 2239 GAGAACTTCTGGAATCGCTAGACAAATTCGGAATGGGATTCAGGATTCAGGATGACAGAAACA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy 2299 CACGATATATAGTGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 15

US-10-718-628-3
; Sequence 3, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10/718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; PRIOR FILING DATE: 1999-10-29

Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
Search completed: February 10, 2006, 04:06:10
Job time : 456.5 secs

```

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-3

Alignment Scores:
Pred. No.: 4.33e-112 Length: 253
Score: 1222.00 Matches: 231
Percent Similarity: 93.7% Conservatives: 6
Best Local Similarity: 91.3% Mismatches: 16
Query Match: 22.1% Indels: 0
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-3 (1-253)

Qy 1579 GTGAATCGACATAATAATGACTAGAGAGAAAGAAATGATGTTTCATGAAATTAAAG 1638
Db 1 MetLysGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGATATTTGGTAATATGGGATGATGTTAAGGCTATTGGTGTGTTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValLeuSerThrGlu 60
Qy 1759 GAAGCAGAGTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
Db 61 GlyValGluPheSerTyrGluTyrThrGlyGluTyrLysAlaGluValAsnPheTyr 80
Qy 1819 AGCGAGAGATTTCTACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTyrProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTTTCCGATTTATGATTCAGTGGATGATCTTAGAGAAAGTGTAT 1938
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120
Qy 1939 CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTGATATGAGGCAATGGCTAATATTCGTGTGCAAGGACCGACAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACATCTTGACTGACTAGCAGTAGCAATGGCAGTGCATGTTGATTTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
Qy 2179 CTTCCTTCAGGTTATGACCATCTGTGCCAGTTCGTAATGCTGTGTCACACTTCCGACTCT 2238
Db 201 LeuProGlyTyrValGlnLeuCysGlnLeuValMetSerGlyGlnLeuSerAspPro 220
Qy 2239 GAGAACTTCTCGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGCAGACAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTyrPasnGlyValGlnGluTyrAlaGluArg 240
Qy 2299 CACGGATATATAGTGGATGTGTCAAAACGCGATACCATTT 2337
Db 2337
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2006, 03:51:42 ; Search time 4 Seconds
(without alignments)

2070.736 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggttacgagc.....atcggtatgatgtcgatc 3156

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 194028

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abse/ABSSWEB_spool/US10681086/runat_09022006_140529_12034/app_query.fasta.1
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abse02p
-USER=US10681086 @CGN_1_1_7 @runat_09022006_140529_12034 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELEX=6 -DELEXT=7

Database :

Published Applications AA New:

1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.ppe:*
2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.ppe:*
3: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.ppe:*
4: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.ppe:*
5: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.ppe:*
6: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.ppe:*
7: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.ppe:*
8: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.ppe:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	828	15.0	332	6	US-10-793-626-540
2	828	15.0	332	6	US-10-793-626-2562
3	548	9.9	381	7	US-11-098-686-10666
4	545.5	9.9	358	6	US-10-467-657-6970
5	405.5	7.4	322	7	US-11-098-686-10665
6	403	7.3	377	6	US-10-793-626-2926
7	373.5	6.8	334	7	US-11-055-822-644
8	350	6.3	416	6	US-10-467-657-5060
9	115	2.1	339	6	US-10-467-657-4318

10	111.5	2.0	401	6	US-10-467-657-1860	Sequence 1860, Ap
11	110	2.0	588	7	US-11-186-284-213	Sequence 213, App
12	110	2.0	596	6	US-10-821-234-1068	Sequence 1068, Ap
13	104	1.9	425	6	US-10-525-710-22	Sequence 22, Appl
14	103.5	1.9	5935	6	US-10-595-561-776	Sequence 776, App
15	102	1.8	348	7	US-11-055-822-662	Sequence 662, App
16	102	1.8	426	6	US-10-525-710-28	Sequence 28, Appl
17	100.5	1.8	464	7	US-11-057-012-41	Sequence 41, Appl
18	98	1.8	4655	6	US-10-595-561-556	Sequence 556, App
19	96.5	1.7	391	7	US-11-055-822-60	Sequence 60, Appl
20	96.5	1.7	391	7	US-11-055-822-300	Sequence 300, App
21	96.5	1.7	391	7	US-11-055-822-328	Sequence 328, App
22	96	1.7	386	6	US-10-467-657-6564	Sequence 6564, Ap
23	95.5	1.7	417	6	US-10-858-730-227	Sequence 227, App
24	95	1.7	950	7	US-11-077-550-118	Sequence 118, App
25	94.5	1.7	428	6	US-10-525-710-20	Sequence 20, Appl
26	94.5	1.7	430	6	US-10-525-710-30	Sequence 30, Appl
27	94	1.7	1216	7	US-11-115-639-44	Sequence 44, Appl
28	94	1.7	1216	7	US-11-115-639-45	Sequence 45, Appl
29	92.5	1.7	423	6	US-10-525-710-44	Sequence 44, Appl
30	92.5	1.7	5024	6	US-10-793-626-2964	Sequence 2964, Ap
31	92	1.7	437	6	US-10-858-730-214	Sequence 214, App
32	92	1.7	437	6	US-10-858-730-283	Sequence 283, App
33	91.5	1.7	4544	7	US-11-076-427A-32	Sequence 32, Appl
34	91	1.6	423	7	US-11-055-822-640	Sequence 640, App
35	90.5	1.6	338	6	US-10-467-657-4976	Sequence 4976, Ap
36	90.5	1.6	345	7	US-11-124-368A-285	Sequence 285, App
37	90.5	1.6	514	6	US-10-840-688-8	Sequence 8, Appli
38	90.5	1.6	907	7	US-11-124-368A-288	Sequence 288, App
39	89.5	1.6	805	6	US-10-467-657-8246	Sequence 8246, Ap
40	89.5	1.6	912	7	US-11-077-550-116	Sequence 116, App
41	89.5	1.6	1706	7	US-11-052-554A-176	Sequence 176, App
42	89	1.6	437	6	US-10-858-730-288	Sequence 288, App
43	89	1.6	914	7	US-11-077-550-120	Sequence 120, App
44	88.5	1.6	944	7	US-11-077-550-122	Sequence 122, App
45	88.5	1.6	970	6	US-10-467-657-5524	Sequence 5524, Ap

ALIGNMENTS

RESULT 1
US-10-793-626-540
; Sequence 540, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 540
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-540

Alignment Scores:
Pred. No.: 2.49e-67 Length: 332
Score: 828.00 Matches: 158
Percent Similarity: 75.0% Conservative: 49
Best Local Similarity: 57.2% Mismatches: 51
Query Match: 15.0% Indels: 18
DB: 6 Gaps: 2

US-10-681-086-1 (1-3156) x US-10-793-626-540 (1-332)

Qy	2307	TATAGTCGATGTGTC	CAAAACGCATACAT	TTTGAATTCGAAGCGCGGAT	-----	2357
Db	57	TyGlyLysValLysLeu	AsnMetIleLeu	AsnAlaLysSerGlyIleCys	AlaGlu	76
Qy	2358	-----	-----	-----	-GAGTCTTACCGGATGGTG	2375
Db	77	AspCysGlyTyrCys	GlyGlnSerValLys	MetLysGlyGlnArgTyrAla	LeuVal	96
Qy	2376	AATAAGGAAACGCTG	TTGAAGCGCGAAGCGGCG	CACGATCTGTAATATCGGCACATAT	2435	
Db	97	GluGlnAspGlnIleLys	GluGlyAlaGlnValAla	ThrGluAsnGlnIleGlyThrTyr	116	
Qy	2436	TGTATCTGCGCAAGCG	CAGAGTCCGTCTAAC	GAGAGAAGTGGATCAGGTCGTAGATCGC	2495	
Db	117	CysIleValMetSerGly	ArgGlyProSerAsn	ArgGluValAspHisIleCysGlyThr	136	
Qy	2496	GTTCAGGAAATTAAG	AAGACGTAT--	GGACTGAAGATTTGTGCATGCTCTGGACATGTTG	2552	
Db	137	ValGluAspIleLysLys	IleHisProGlnLeu	ValIleCysAlaCysLeuGlyLeuThr	156	
Qy	2553	AAGCCAGCAGGCGCA	AGCGCTCAAGATGCG	AGGAGTAGACCGCTATAATCATATTTG	2612	
Db	157	LysGluGluGlnAlaLys	LeuLysLeuLysAla	AlaGlyValAspArgTyrAsnHisAsnLeu	176	
Qy	2613	AATACGTCACAGAGAA	ACCATTCAAAACATC	CACACCTCACATACATACGATCAGCAGTCC	2672	
Db	177	AsnThrSerGluArgTyr	HisAspGluValThr	HisThrHisThrTyrGluAspArgVal	196	
Qy	2673	AATACGGTTCGAAAT	CGCAAAAGATCGGGG	CTGTCTCCGTGTTCAGGCGCATATTCGGG	2732	
Db	197	AsnThrValGluMetLys	AspAsnAsnIleSerProCysSerGlyValIleCysGly	216		
Qy	2733	ATGAGGAGAGAGAAAC	GAGATGTCATTGCAT	CGCCAAAAGCTTGAAGCTCTTGACGCG	2792	
Db	217	MetGlyGluSerAsnGlu	AspIleIleAspMetAla	PheAlaLeuArgAlaIleAspAla	236	
Qy	2793	GATTCCATTCTGTGA	ATTTTTCATGTCAAAT	TGATGGCAGCGCTTTAGAAGCGCTCAAC	2852	
Db	237	AspSerIleProIleAsn	PheLeuHisProIleLysGlyThrLysPheGlyGlyLeuAsp	256		
Qy	2853	GAATTAACCCGCTGT	ATGTTTAAAAAGTCTG	CGCGCTGTTCGCTTTTATCAATCCATCA	2912	
Db	257	LeuLeuSerProMetLys	CysLeuArgIleIleAlaMetPheArgLeuIleAsnProThr	276		
Qy	2913	AAAGAAATTCGCATT	TCGGGAGGAGAGAGGT	CAATCTCCGACATTGCGACCATTAGGG	2972	
Db	277	LysGluIleArgIleAla	GlyArgGluValAsnLeu	ArgSerLeuGlnProLeuAla	296	
Qy	2973	CTTTTACCGCCCAACT	CCATTTTGTCCGAGACT	TACTTAACAACCTCGCGCCCAAGAGGAG	3032	
Db	297	LeuLysAlaAlaAsnSer	IlePheValGlyAspTyrLeuIleThrGlyGlyGlnProAsn	316		
Qy	3033	ACGGAGGATCATAAAT	CTGCTGAGTGATTTAG	CGCTTTGAAGTTGAATCA	3080	
Db	317	GluGluAspTyrArgMet	IleGluAspLeuGlyPheGluIleAspSer	332		

RESULT 2

```

US-10-793-626-2562
; Sequence 2562, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,358
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2562
; LENGTH: 332
; TYPE: PRT

```

```
US-11-098-686-10666
; Sequence 10666, Application US/11098686
; Publication No. US2006024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10666
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10666

Alignment Scores:
Pred. No.: 7.32e-42 Length: 381
Score: 548.00 Matches: 131
Percent Similarity: 56.9% Conservative: 54
Best Local Similarity: 40.3% Mismatches: 132
Query Match: 9.9% Indels: 8
DB: 7 Gaps: 3

US-10-681-086-1 (1-3156) x US-11-098-686-10666 (1-381)
QY 151 TGGTCTCAACAAATTATTAGGCTCGCAAGCATAGAGCTTTGATCGATGAGCCCAA 210
DB 35 PheSerThrAsnAspTyrLeuGlnLeuAlaTyrHisProAsnLeuLeuAsnAlaIle 54
QY 211 ACAGCATTCGACAAATTGGGACAGGAGCGGTTTACGTTTAAAGCAGCAGCAATTCG 270
DB 55 MetThrGlySerThrTyrGlyValGlySerThrGlySerArgLeuLeuSerGlyAsnAsn 74
QY 271 GTCTGGCATGAAAGCTAGAAAAGAGATTGCCAGCTTTAACTGACAGAGCGGCTCG 330
DB 75 GluLeuPheGluArgLeuGluThrThrIleAlaGlnAspLysHisThrGluThrAlaMet 94
QY 331 CTGTTTTCGAGCGTTACTTGGCCAAATGTCGGTGTCTTTCATCTTCCCTTGCAGAA- 384
DB 95 LeuPheIleSerGlyPheGlnThrAsnValSerValLeuSerAlaLeuLeuAspHisHis 114
QY 385 -----AAGGAAGATGTCATTTTAAGTGACAGCTCAATCATGCAAGTATGATCGAC 435
DB 115 ValLeuLysMetGlnProLeuValPhePheAspLysLeuAsnHisAlaSerLeuTyrGln 134
QY 436 GGCTGCCGACTTCTTAAGGCTGATACAGTTGTTTATCGGCATATTTGATATGATGATCTT 495
DB 135 AlaValPheLeuSerLysAlaGluLeuLeuArgTyrTyrHisAsnAsnMetGluHisLeu 154
QY 496 GAAACAAGCTGAATGAAACACAGCGTTATCAGCGC---CGTTTATCGTAAACAGACGGA 552
DB 155 SerSerLeuLeuLysLysTyrLysAspAsnArgProLysPheIleValThrGluThr 174
QY 553 GTATTCAGCATGGATGGCAACATCGCCCTTTCATGATCAGATCATCTCATTGCGAAACGC 612
DB 175 LeuPheGlyMetAspGlyAspIleAlaProLeuThrAspIleValSerLeuSerSerGln 194
QY 613 TATCATCCTTCTGTCGTCGTGATGATGCCAGCAACAGAGATTTCGGCGGATTCGGGA 672
DB 195 TyrGlnThrPheLeuTyrLeuAspGluAlaHisAlaThrGlyLeuThrGlyIleHisGly 214
QY 673 CAAGGACAGAGTGAATACTTTGGTGT------TGTCGCCACATTTGTTATCGCACCTTA 726
DB 215 TyrGlyLeuSerThrThrValAsnLeuSerHisIleProHisIleValMetGlyThrPhe 234
QY 727 AGCAAGCTGTTGGCGCGGAGGAGGTTTTCGCGCAGGATCAGCGGCTCTTCATCGACTTT 786

US-10-467-657-6970
; Sequence 6970, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6970
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6970

Alignment Scores:
Pred. No.: 1.2e-41 Length: 358
Score: 545.50 Matches: 133
Percent Similarity: 58.0% Conservative: 63
Best Local Similarity: 39.3% Mismatches: 127
Query Match: 9.9% Indels: 15
DB: 6 Gaps: 7

US-10-681-086-1 (1-3156) x US-10-467-657-6970 (1-358)
QY 121 AGGAATATTGATGGCGAAATCAACGGTCT-----TGGTCTCAACAAATTATTAGGG 174
DB 6 ArgTyrIleThrArgGluAsnArgLysMetLeuAsnMetSerSerAsnAspTyrLeuGly 25
QY 175 CTCGCAAGCGATAGAGCTTTGATGTCGATGCGACCCCAACAGCATTCGACGAAATTTGGACA 234
DB 26 LeuAlaSerAspGluAsnLeu-----ArgArgSerPheLeuGlnGlnTyrGlyGly 42
QY 235 GGA-----AGCAGCGGTTTCACGTTTAAACGACGCAATTCGGTCTCGCAT 279
DB 43 AsnPheProSerPheThrSerSerArgLeuLeuThrGlyAsnPheProIleTyr 62
```

```
QY 280 GAAAGCTGAGAAAGAGATTGCCAGCTTTAAACTGACAGAAAGCGGCCTCTGTTTCG 339
Db      |||||:::  ::|||  |||||:::
QY 63 ThrAspLeuGluLeuValAlaGlnArgPheGlnArgGluSerAlaLeuLeuPheAsn 82
Db      |||||:::  ::|||  |||||:::
QY 340 AGCGTTACTTGGCAATGTCGGTGTCTTTCATCTCTTCCAGAAAGAGAGATGTCATT 399
Db      |||||:::  ::|||  |||||:::
QY 83 SerGlyTyrHisAlaAsnLeuGlyLeuProAlaLeuThrThrLysSerLeuLeu 102
Db      |||||:::  ::|||  |||||:::
QY 400 TTAAGTGACCAAGCTCAATCATGCAAGATGATGACCGCTCCGACTTCTTAAGGTCAT 459
Db      |||||:::  ::|||  |||||:::
QY 103 LeuAlaAspLysPheValHisAlaSerMetIleAspGlyIleArgLeuSerArgCysAla 122
Db      |||||:::  ::|||  |||||:::
QY 460 ACAGTTGTTTATCGCATATTGATGAATGATCTTGAACAAAGCTGAATGAA---ACA 516
Db      |||||:::  ::|||  |||||:::
QY 123 PhePheArgTyrArgHisAsnAspTyrGluHisLeuLysAsnLeuLeuGluLysAsnVal 142
Db      |||||:::  ::|||  |||||:::
QY 517 CAGCGTTATACGCCCGCTTTTATCGTAAACAGACGGAGTATTACAGATGGATGGCACAAATC 576
Db      |||||:::  ::|||  |||||:::
QY 143 GlyLysPheAspArgThrPheIleValThrGluSerValPheSerMetAspGlyAspVal 162
Db      |||||:::  ::|||  |||||:::
QY 577 GCCCTCTTGATCAGATCATCTCACTTCGGAACGCTAT---CATGCCCTTCGTGTCGT 633
Db      |||||:::  ::|||  |||||:::
QY 163 AlaAspLeuLysGlnLeuValGlnLeuLysLysGlnPheProAsnThrTyrLeuTyrVal 182
Db      |||||:::  ::|||  |||||:::
QY 634 GATGATGCCCAAGCAAGAGGATTGTCGGGATTCGGGACCAAGCAACGAGTGAATACTTT 693
Db      |||||:::  ::|||  |||||:::
QY 183 AspGluAlaHisAlaIleGlyValTyrGlyGlnAsnGlyLeuGlyIleAlaGluArgAsp 202
Db      |||||:::  ::|||  |||||:::
QY 694 GGTGTTTGTCCC-----GACATTGTTATCGCACCTTTAAGCAAAAGCTGTTGGCCCGGAA 747
Db      |||||:::  ::|||  |||||:::
QY 203 AsnLeuIleAlaGluIleAspLeuLeuValGlyThrPheGlyLysAlaLeuAlaSerVal 222
Db      |||||:::  ::|||  |||||:::
QY 748 GGAGGTTTGGCGGAGGATCAGCGGTCTTCATGACATTTTTCGTGAACCATGCCAGAAC 807
Db      |||||:::  ::|||  |||||:::
QY 223 GlyAlaTyrAlaValCysAsnGlnValLeuLysGluCysLeuIleAsnGlnMetArgPro 242
Db      |||||:::  ::|||  |||||:::
QY 808 TTTATCTTTCAAAACCGCTATTCCGCCAGCAGCTGTGCGGTGCTCAGCAGGCTTTCAAC 867
Db      |||||:::  ::|||  |||||:::
QY 243 LeuIlePheSerThrAlaLeuProPheAsnValAlaTyrThrTyrPheIlePheGlu 262
Db      |||||:::  ::|||  |||||:::
QY 868 ATCATTTGAAGCCAGCAGGGGAAACGACAGCTTTTATTTTCTTATATCATGATCATG 927
Db      |||||:::  ::|||  |||||:::
QY 263 ArgLeuProGlnPheSerLysGluArgSerHisLeuGluGlnLeuSerAlaPheLeuArg 282
Db      |||||:::  ::|||  |||||:::
QY 928 ACCAGCTGAGAGATATCGGTATGCTGTGTAAGAGAGATCACACCGCATTTTCTCTGA 987
Db      |||||:::  ::|||  |||||:::
QY 283 ArgGluValAlaHisArgThrGlnIleMetProSerGlu---ThrCysIleValProTyr 301
Db      |||||:::  ::|||  |||||:::
QY 988 GTCATTGGCGATGCCCATAAACGGTCTCTATTGCTGAAAAAATGCGAGGGCAAGGAATT 1047
Db      |||||:::  ::|||  |||||:::
QY 302 IleLeuGlyGlyAsnGluAlaThrLeuAlaLysAlaGluTyrLeuGlnGlyGlnGlyTyr 321
Db      |||||:::  ::|||  |||||:::
QY 1048 TATGCTCTGCCATTGCGCCCAACCGTTGCGCGGGTGAAAGCCGGAATTGCA 1101
Db      |||||:::  ::|||  |||||:::
QY 322 TyrCysLeuProIleGlyProProThrValProLysAsnThrSerArgIleArg 339
Db      |||||:::  ::|||  |||||:::
```

RESULT 5

```
US-11-098-686-10665
; Sequence 10665, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 10665
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10665

Alignment Scores:
Pred. No.: 6,08e-29      Length: 322
Score: 405.50           Matches: 90
Percent Similarity: 58.9%      Conservative: 56
Best Local Similarity: 36.3%    Mismatches: 93
Query Match: 7.4%             Indels: 9
DB: 7                       Gaps: 5
```

US-10-681-086-1 (1-3156) x US-11-098-686-10665 (1-322)

```
QY 2346 AAAGCGCCGATGAGTCTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGCGCGGAAG 2405
Db      |||||:::  ::|||  |||||:::
QY 71 LysThrAspLeuGlnLysGluProLeuValAspIleAspThrLeuIleSerGluAlaLys 90
Db      |||||:::  ::|||  |||||:::
QY 2406 CGGGCGCACGATCTGAATATCGGCACATATTGTATCTGCGCAAGCGCGCAGAGTCCGCT 2465
Db      |||||:::  ::|||  |||||:::
QY 91 LysAlaLysGluLeuGlySerThrArgPheCysMetGlyAlaAlaTrpArgSerProLeu 110
Db      |||||:::  ::|||  |||||:::
QY 2466 AACAGAGAAGTGCATCAGTCTGATGCGGTTTCAGGAAATTTAAAGAGACGATATGGACTG 2525
Db      |||||:::  ::|||  |||||:::
QY 111 AspLysAspLeuLysIleValCysGlnMetIleGluGluValLysLys---LeuGlyLeu 129
Db      |||||:::  ::|||  |||||:::
QY 2526 AAGATTTGTCATGTCCTTGGACTGTTGAAGCCAGAGCAGCGCAAGCGCTCAAGATGCA 2585
Db      |||||:::  ::|||  |||||:::
QY 130 GluThrCysValThrLeuGlyPheLeuLysGluHisGlnIleAlaMetLeuLysLysAla 149
Db      |||||:::  ::|||  |||||:::
QY 2586 GGAGTAGACCGCTATATCATTAATCACTGCACAGAGAACACCATCTCAACATCACA 2645
Db      |||||:::  ::|||  |||||:::
QY 150 GlyLeuAspPheTyrAsnHisAsnMetAsnThrSerProGluPheTyrGluHisIleAla 169
Db      |||||:::  ::|||  |||||:::
QY 2646 ACCTCACATACATACGATGACAGATCAATACGCTTGAATTCGCAAAAGAAATCGGGCTG 2705
Db      |||||:::  ::|||  |||||:::
QY 170 ThrThrHisThrPheAspAspArgLeuAlaThrLeuLysAlaValArgLysPheGlyIle 189
Db      |||||:::  ::|||  |||||:::
QY 2706 TCTCCGTTTCAGCGCCATTTATCGGATGAAGAGACGAAACAGAGATGTCATTGACATC 2765
Db      |||||:::  ::|||  |||||:::
QY 190 LysLeuCysSerGlyGlyIleIleGlyLeuGlyGluThrIleAspAspArgIleSerMet 209
Db      |||||:::  ::|||  |||||:::
QY 2766 GCCAAAAGCTGAAGGCTCTTGAC-----GCGATTCCTATTCCTGTGAAATTTTTCAT 2819
Db      |||||:::  ::|||  |||||:::
QY 210 LeuLeuLeuLeuAlaThrLeuGluGluGlnProGluSerValProIleAsnArgPheVal 229
Db      |||||:::  ::|||  |||||:::
QY 2820 GCAATTGATGCGCAGCGCTTAGAAGCGCTCAACGAATTAACCCGCTGTATTGTTTAAAA 2879
Db      |||||:::  ::|||  |||||:::
QY 230 LysValAlaGlyThrProLeuAsnProGlnSerAspIleAspProPheAspPheValArg 249
Db      |||||:::  ::|||  |||||:::
QY 2880 GTGCTGCGCTGTTCCGTTTATCAATCAATCAAAAGAAATTCGCATTTCGCGAGGAAGA 2939
Db      |||||:::  ::|||  |||||:::
QY 250 IleIleAlaLeuThrArgIleLeuMetProLysSerTyrIleArgLeuAlaAlaGlyArg 269
Db      |||||:::  ::|||  |||||:::
QY 2940 GAGTCTC-----AATCTCCGCACATTGCGCCATTAGCGCTTTACGCCGCAAACTCC 2990
Db      |||||:::  ::|||  |||||:::
QY 270 GluGlnMetSerAspGluLeuGlnThrLeuCysPheMetGly-----GlyAlaAsnSer 287
Db      |||||:::  ::|||  |||||:::
QY 2991 ATTTTTCGCGAGAC---TACTTAAACAACTGCCGCGCAAGAGAGACGCGGATCATAAA 3047
Db      |||||:::  ::|||  |||||:::
QY 288 IlePheTyrGlyArgLeuLeuThrThrAspGlyProGlnProGluGlnAspLeu 307
Db      |||||:::  ::|||  |||||:::
QY 3048 ATGCTGAGTGATTTAGCTTTGAA 3071
Db      |||||:::  ::|||  |||||:::
QY 308 LeuPheGlnLysIleGlyLeuGlu 315
Db      |||||:::  ::|||  |||||:::
```

RESULT 6

```
US-10-793-626-2926
; Sequence 2926, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
```

```
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2926
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2926

Alignment Scores:
Pred. No.: 1,09e-28 Length: 377
Score: 403.00 Matches: 114
Percent Similarity: 51.1% Conservative: 76
Best Local Similarity: 30.6% Mismatches: 138
Query Match: 7.3% Indels: 44
DB: 6 Gaps: 10

US-10-681-086-1 (1-3156) x US-10-793-626-2926 (1-377)

Qy 40 TTAACGAGCGGTAGACAGATGAAGACCGCGGTACATCTAACCCTCGGTCATG 99
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 IleLysAlaGlnLeuLysGlnIleGlnAspLysGlyLeuTyrArgGluLeuGlnProile 22

Qy 100 GATGAGCGCGGTTCAGAGAGCAATATTCATGCGCAAAATCAACCGTCTGCTCTCA 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 23 GlnSerValGluLysGlnTyrIleTyrIleAsnAspGlnSerTyrIleAsnPheThSer 42

Qy 160 AACAAATTATTAGCGCTCGCA-----AGCATAGACGCTTTGATCGATCGACGCC 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 43 AsnAspTyrLeuGlyIleGlyGlnValGluTyrGlnProGlnAsnPheLeuAspPheile 62

Qy 208 CAACAGCATTCAGCAATTTGGACAGGACGCGGTTCACGTTTAAACGACAGGCAAT 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 LysThr-----TyrSerIleHisLeuSerSerSerArgLeuValSerGlyAsn 78

Qy 268 TCGGTCTGGCATGAAAGCTAGAAAGAGATTCGCAGCTTTAAACTGACAGAAAGCGCC 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 79 SerValValTyrGlnGlnLeuGluGlnIleSerGluHisPheAsnPheGluAspAla 98

Qy 328 CTGCTGTTTCGACGGTTACTTGGCCAAATGTCGGTGTCTCTTCATCTCTGCGCAAGAA 387
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 LeuilePheAsnSerGlyTyrAspAlaAsnLeuAlaValPheAsnIlePheLysAsnAsn 118

Qy 388 GAAGATGCTATTTAAGTGACCACTCAATCATCAAGTATGATCGAGCGGTGCGGACTT 447
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 AsnIleValIlePheSerAspGlnGlnAsnHisAlaSerIleAspGlyIleLysLeu 138

Qy 448 TCAAGGCTGATCAGATGTTTATCGGCATATTCATATGATGATCTTGAACCAAGCTG 507
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 139 SerGlyLeuSerLysValIleTyrGlnHisLeuAsnTyrAspAspLeuGluSerHisLeu 158

Qy 508 ---AATGAACACACAGCGTTATCAGCGCGTTTATCGTAACACAGCGGAGTATTCAGCATG 564
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 AlaArgHisThrAsnProAspValGlnLysValIleValSerAspSerValPheSerThr 178

Qy 565 GATGCGCAATCGCCCTCTTGATCAGATCATCTCATTGCGGAACGCTATCGCTTC 624
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 AsnGlyThrLysAlaAspIleAsnArgLeuValHisLeuLysGlnArgTyrAsnAlaile 198

Qy 625 GTGTCGTGTTGATGCCACGACAGAGAGTTTGGCGGATTCGGGCAACGAAGCAGT 684
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 LeuileleAspAlaSerHisSerLeuGlyLeu-----AsnLeuPhe 212

Qy 685 GAATACTTGTGTTGTTGTCGCGACATTTGTTATCGGCACCTTAAGCAAAAGCTGTGGCGCG 744
```

```
Db 213 GluTyrHisAlaAspIle---AspIleValThrSerSerLeuSerLysAlaTrpGlyAla 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 745 GAAGGAGGTTTTCGGCGAGGATCAGCGTCTTCATCGACTTTTTCGCTGAACCATGCCAGA 804
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 HisGlyGlyValIlePheSerSerLysAspIleLysAspLeuIleAsnLysGlyArg 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 805 ACATTATCTTCAACCGCTATTCCGCCAGCCAGCTGTGCGGTGCTCAGCAGGCTTTC 864
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 252 SerLeuileTyrSerSerLeuPro-----SerTyrHisLeuTyrPhe 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 865 -----ACATCATTTGAAGCCAGC-----AGGAAAAACGACGCTT 900
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 267 IleGlnValSerLeuGlnHisValIleGluAspThrTyrArgArgGluLysLeuAsnVal 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 901 TTATTTTCTTATATCAGC-----ATGATCAGACACAGCTCTGAAGAAT 942
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 287 LeuSerGluTyrPheAsnHisGlnPheMetGluLeuPheProAspGlnProLeuSerAsn 306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 943 ATGGGTTTATGTGTGAAGGAGATCACACACCGGATTATTCCTGTAGTCATTGGCGATGCC 1002
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 -----ThrProIleLysAsnIleValCysAspSerLeu 317
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1003 CATAAACGGTCTTATTTGCTGAAAAAAGCTGACGGCAAGGAATTTATGCTCTGCCATT 1062
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 318 AlaSerAlaGlnAlaGlnTyrAspMetLeuPheGluHisGlyIlePheValSerTyrLeu 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1063 CGCGCCGCAACCGTTGCGCGCGGTGAAGCGCGATT 1098
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 338 ArgTyrProThrValSerGlnLeuThrLeuArgile 349

RESULT 7
US-11-055-822-644
; Sequence 644, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberkauer, Gregor
; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 644
; TYPE: PRT
; LENGTH: 334
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-644
```



```

Db 295 AspProGlnProGluSerValProIleAsnArgLeuValLysValGluGlyThrProLeu 314
Qy 2841 GAAGCGCTCAACGATTAACCGCTGATTTGTTTAAAGTGGCGCTGTCCTGTTT 2900
Db 315 AlaAspAlaGluAspLeuAspTrpThrGluPheValArgThrValSerValAlaArgIle 334
Qy 2901 ATCAATCCATCAAAAGAAATTCGATTTCCGAGAGAGAGAGAGTCAATCTCCGACATTTG 2960
Db 335 ThrMetProGlnSerTyrValArgLeuSerAlaGlyArgSerAsnMetProGluAlaMet 354
Qy 2961 CAGCAATTAGGCTTTAGCCG---GCAAACTCCATTTTGTTCGAGACTACTTAACACT 3017
Db 355 GlnAlaMetCysPheMetAlaGlyAlaAsnSerIlePheTyrGlyAspLysLeuLeuThr 374
Qy 3018 GCCGGGCAAGAGAGAGAGAGAGATCAATAAATGCTGAGTGTAT----- 3059
Db 375 ThrGlyAsnProAspGluAspGlyAspArgIleLeuMetGluLysLeuAsnLeuTyrPro 394
Qy 3060 TTAGGCTTTGAAGTTGAATCA-----GTCGAAGAAATGAAGGCTAGT 3101
Db 395 LeuGlnPheGluProGluGlyGluValAlaGluValGluLysAlaSer 410

RESULT 9
US-10-467-657-4318
; Sequence 4318, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: Seqwin99, version 1.04
; SEQ ID NO 4318
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4318

Alignment Scores:
Pred. No.: 0.0155 Length: 339
Score: 115.00 Matches: 64
Percent Similarity: 39.1% Conservative: 56
Best Local Similarity: 20.8% Mismatches: 103
Query Match: 2.1% Indels: 84
DB: 6 Gaps: 17

US-10-681-086-1 (1-3156) x US-10-467-657-4318 (1-339)
Qy 265 AATTCGGTCTGCAT-----GAAAGCTAGAAAG 294
Db 2 HisLysLeuThrHisIleSerAsnIleTyrThrArgProAlaGlnGluLeuAlaGln 21
Qy 295 AAGATTGCCACTTTAACTGACAGAACGCCCTGCTGTTTCGAGCGGTACTTGCC 354
Db 22 LysLeuValLysHisSerPheAlaAspLysValPhePheCysAsnSerGlyAlaGluAla 41
Qy 355 AATGTCGGTCTCTTCATCTCTGCCAGAA-----AAG 387
Db 42 AsnGluAlaAlaLeuLysLeuAlaArgLysTyrAlaArgAspArgPheGlyGlyLys 61
Qy 388 GAAGATGTCATTTTAAAGTACCAGCTCAATCATGCAAGTATGATC----- 432
Db 62 SerGluIleValAlaCysIleAsnSerPheHisGlyArgThrLeuPheThrValSerVal 81

```

```

Qy 433 -----GACGGCTGCCGACTTCTTAAGGCTGATACAGTTGTT 468
Db 82 GlyGlyGlnProLysTyrSerLysAspTyrAlaProLeuProGlnGlyIleThr----- 99
Qy 469 TATCGGCATATTGATATGAATGAT-----CTTCAAAAACAGCTGAATGAACACAG 519
Db 100 -----HisValProPheAsnAspIleAlaAlaLeuGluAlaValGlyGlu----- 115
Qy 520 CGTTATCAGCGCCGTTTTATCGTAACAGAGGAGTATTACAGCATGGATGCACAAATGCC 579
Db 116 -----GlnThrCysAlaValIleLeuProIleGlnGlyGluSerGlyIleLeuPro 133
Qy 580 CCTCTTCATCAGATCATCTCCTGCGAAA-----CGTATCATGCTCTCGTG 627
Db 134 AlaThrAlaGluTyrLeuGlnAlaAlaArgLeuCysAspArgHisAsnAlaLeuLeu 153
Qy 628 GTCCTTCATGATCCCGCAACAGGAGTTTTGGCGGATTCGGGACAA---GGAACGAGT 684
Db 154 IleLeuAspGluValGln---ThrGly---MetGlyHisThrGlyArgLeuPheAlaTyr 171
Qy 685 GAATACCTTTGGTGTTCGCGACATTTGTTATCGGCACCTTTAAGCAAGCTGTTGGCGCG 744
Db 172 GluHisTyrGlyValValProAspIleLeu-----SerSerAlaLysAlaLeuGlyCys 189
Qy 745 GAAGGAGCTTTTGGCGGAGGATCAGCGCTCTTCATCGACTTTTCTGCTG----- 792
Db 190 -----GlyPheProIleGlyThrMetLeuAlaThrGluLysIleAlaAlaPheGln 207
Qy 793 -----AACCATGCCAGAACATTTATCTTCAAAACCGCTATTCCCGCAGCCAGCTGTGCG 846
Db 208 ProGlyThrHisGlySerThrPhe-----GlyGlyAsnProMetAlaCysAla 223
Qy 847 GCTGCTCAGCAGGCTTCAACATCATTAAGCCAGCAGGAGGAAAAACGACAGCTTTTATT 906
Db 224 ValGlySerArgAlaPheAspIleAlaAsnAlaPro---GluThrLeuHisAsnValArg 242
Qy 907 TCTTATATCAGCATGATCAGAACCATCTGTAAGAAATATGGTTATGTGCTGAAAGGAGAT 966
Db 243 SerGlnGlyGlnLysLeuGlnThrAlaLeuAspLeuGly----- 256
Qy 967 CACACACGATTTATCTCTGCTAGTCTATGCGGATGCCCATAAACGGTCTATTGTCTGAA 1026
Db 257 -----ArgLysThrGlyLeuPheSerGln 264

Qy 1027 AAATCGCAGGCGCAGGGAATT 1047
Db 265 ---ValArgGlyMetGlyLeu 270

RESULT 10
US-10-467-657-1860
; Sequence 1860, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: Seqwin99, version 1.04
; SEQ ID NO 1860
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1860

Alignment Scores:

```



```
Pred. No.: 0.0346 Length: 401
Score: 111.50 Matches: 60
Percent Similarity: 41.5% Conservative: 38
Best Local Similarity: 25.4% Mismatches: 78
Query Match: 2.0% Indels: 60
DB: 6 Gaps: 11

US-10-681-086-1 (1-3156) x US-10-467-657-1860 (1-401)
QY 90 GCGGTCAATGATGAGCGCGGTTCAGAGAGGAATATTGATGCGGAAATCAACCGT 149
Db 15 AlaLeuSerProThrPheThrArgGluGlu-AlaAspAlaValSerLysVal 34
QY 150 GTGTCCTCAAA--AAATTATTAGGCGTCGCAAGCATAGACGTTTGATCGATGACG 206
Db 34 uLeuSerAsnLysValAsnTyr----- 41
QY 207 CCAAACAGCATTCGACGAATTTGGACAGAGGAGCGGTTCACGTTTAACGACAGGCAA 266
Db 42 -----TriThrGlySerGlu----- 46
QY 267 TTCGGTCTGGCATGAAGAGTAGAAAGAGATTGCCAGCTTTAACTGACAGAGCGGC 326
Db 47 -----CysArgGluPheGluGlyPheAlaAlaPheAlaGlyThrArgTyrAl 63
QY 327 CTGCTGTTTTCGAGCGGTACTTGGCCAAATGTCGGTGTCTCTTCATCC-----TTGCC 380
Db 63 aValAlaLeuSerAsnGlyThrLeuAlaLeuAspAlaAlaLeuLysAlaIleGlyIleG1 83
QY 381 AGAAGAGGAGATGTCATTTTAAGTGACGAGCTCAATCATGCAAGTATGATGACGCGGTG 440
Db 83 yAlaGlyAspValIleValThrSerArgThrPheLeuAlaSerAla-----SerCy 101
QY 441 CCGACTTCTAAGGCTGATACAGTTGTTTATCGGCATATTGATGATGATGATGTTGAAA 500
Db 101 sileValAsnAlaGlyAlaAsnProValPheAlaAspValAspLeuAsnSer---GlnAs 120
QY 501 CAAGCTGAATGAACACACAGCGTTATCGCGCGGTTTATCGTAACAGACGGA----- 552
Db 120 nileSerAlaGluThrVal-----LysAlaValLeuThrProAsnThrLysAl 136
QY 553 -----GTATTACATGATGGAGCAATCGCCCTCTTGATCATGATCATCTCATCTGC 605
Db 136 aValIleValValHisLeuAlaGlyWetProAlaGluMetAspGlyIleMetAlaLeuAl 156
QY 606 GAAAGCTCATGCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 657
Db 156 aLysGluHisAspLeuThrValIleGluAspCysAlaGlnAlaHisGlyAlaThrTyrLy 176
QY 658 -----TTGGCGGATTCGGGACAGGAACAGATGATGATGATGATGATGATGATGATG 707
Db 176 sGlyLysSerValGlySerIleGlyHisValGlyAlaThrPhe-----CysGlnAs 194
QY 708 CATGTTATCGGACCTTAAGCAAGCTGTTGGCGCGGAAGAGGT 753
Db 194 pLysIleIleThr-----GlyGlyGluGlyGly 204

RESULT 11
US-11-186-284-213
; Sequence 213, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
```

```
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; PRIOR FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 213
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-11-186-284-213

Alignment Scores:
Pred. No.: 0.0554 Length: 588
Score: 110.00 Matches: 113
Percent Similarity: 35.4% Conservative: 78
Best Local Similarity: 20.9% Mismatches: 171
Query Match: 2.0% Indels: 178
DB: 7 Gaps: 30

US-10-681-086-1 (1-3156) x US-11-186-284-213 (1-588)
QY 718 GGCACCTTAAGCAAGCTGTGGCGGAGGAGGTTTTCGGCGA-----GGATCAGCGTC 774
Db 37 GlyLysAspAlaLysThrThrAsnGlyHisGlyGlyGluAlaAlaGluGlyLysSerLeu 56
QY 775 TTCATCGACITTTTGTGTCGACCATGCCAGACATTTTATCTTTCAACCGCTATTCCGCCA 834
Db 57 GlySerAlaLeuLysProGlyGluGlyArgSerAlaLeuPhe----- 70
QY 835 GCCAGCTGTGGCGTGTCTACGAG-----GCTTTCAACATCATTTGAAGCCAGC 882
Db 71 -----AlaGlyAsnGluThrArgArgProIleIleGlnPheValGluSerGly 86
QY 883 AGGGAAGAACGACAGCTTTTATTTCTATATCATGATGATGATGATGATGATGATGATGATG 942
Db 87 AspAspLysAsnSerAsnTyrPheSerMetAspSerMet----- 99
QY 943 ATGGGTTATGTCGTGAAGAGATCACACCGGATTTATCTCTGTAGTCATTTGGCGATGCC 1002
Db 100 -----GluGlyLysArgSerProTyrAlaGlyLeuGlnLeuGlyAlaAla 114
QY 1003 CATAAA---ACGGTCTCTATTGCTGAAAAAATGTCAG---GGCAAGGGAATTTATGCTCCT 1056
Db 115 LysLysProProValThrPheAlaGluLysGlyGlyLeuArgLysSerIlePheSerGlu 134
QY 1057 GCCATTGCGCGCAACCGTTGCG-----CCGGTGAAGCCGGAATTCGAAGC--- 1104
Db 135 Ser---ArgLysProThrValSerIleMetGluProGlyGluThrArgArgAsnSerTyr 153
QY 1105 -----TTGGCGAGCGGTCGAGATCAGGGAATGAGTTTATA----- 1140
Db 154 ProArgAlaAspThrGlyLeuPheSerArgSerLysSerGlySerGluValLeuValCys 173
QY 1141 -----AAATAAAAAACACCTGAAAAAGGTCCTCTTTTGTGATGTTTGTG 1185
Db 174 AspSerCysIleGlyAsnLysGlnLysAlaValLysSerCysLeuValCysGlnAlaSer 193
QY 1186 AACTGTTCTTCTTATCTTATGATATAGAAATAAGCTCATTTTATTTATTTTAG 1245
Db 194 PheCysGluLeuHisLeuLysProHisLeuLysGlyAlaAlaPheArgAsp---HisGln 212
QY 1246 TTGCTGAAA-----GGTCCGTTGAGTGTGTTGTTGTTGTTGTTTAAAG 1290
Db 213 LeuLeuGluProIleArgAspPheGluAlaArgLysCys-----ProVal 227
```



```
Db 254 -----|||||:
; LeuCyMetPheGlnGluHis 260
Qy 1398 -----
Db 261 LysAsnHisSerThrValThrValGluGluAlaLysAlaGluLysGluThrGluLeuSer 280
Qy 1399 -----CTAATAGTAGCATTATTACAGATGAAATAATCAAGGTTTAGTGG 1443
; :|||:|||||:
Db 281 LeuGlnLysGluGlnLeuGlnLeuLysIleGluIleGluAspGluAlaGluLysTrp 300
Qy 1444 ACAAGACAAAAGTGAAGTGAAGCACCATTGTGCTTAGAAGACAGAGTTATTAAATAGCTG 1503
; :|||:|||||:
Db 301 GlnLysGluLys-----AspArgIleLysSerPheThrThr----- 312
Qy 1504 AATAGAACGGTCTCTC---CAATATCTTTATTAGAAAAGCAAAATCTAAATATATCT 1560
; :|||:|||||:
Db 313 AsnGluLysAlaIleLeuGluGlnAsnPheArgAspLeuValArgAspLeuGluLysGln 332
Qy 1561 GAAAAGGGAATCAGA-----ATAGTGAATGGACCA 1590
; :|||:|||||:
Db 333 LysGluGluValArgAlaLeuGluGlnArgGluGlnAspAlaValAspGlnValLys 352
Qy 1591 ATAATAATG---ACTAGAGAAAGAATGAAGATTGTTCTAGAAATTAAGGAA----- 1641
; :|||:|||||:
Db 353 ValIleMetAspAlaLeuAspGluArgAlaLysValLeuHisGluAspLysGlnThrArg 372
Qy 1642 CGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTT--- 1698
; :|||:|||||:
Db 373 GluGlnLeuHisSerIleSerAspSerValLeuPheLeuGlnGluPheGlyAlaLeuMet 392
Qy 1699 -----GGTGGT--- 1704
; :|||:|||||:
Db 393 SerAsnTyrSerLeuProProLeuProLeuProThrTyrHisValLeuGluGluGly 412
Qy 1705 -----CAGACTGATGGCCCTATTTCGGAT-----ATT 1731
; :|||:|||||:
Db 413 LeuGlyGlnSerLeuGlyAsnPhelysAspAspLeuLeuAsnValCyMetArgHisVal 432
Qy 1732 GAGATGATGTGTGTC-----ATGTCAACAGAGGACGAGATTCAGCCAT----- 1776
; :|||:|||||:
Db 433 GluLysMetCysLysAlaAspLeuSerArgAsnPheIleGluLysArgAsnHisMetGluAsn 452
Qy 1777 -----GAATGGACAACC-----GGTGGTGAAGGTG 1803
; :|||:|||||:
Db 453 GlyGlyAspHisArgTyrValAsnAsnTyrThrAsnSerPheGlyGlyGluTrpSerAla 472
Qy 1804 GAAGTGAATTTTGATAGCGAAGAGATCTACTAGATTATGCATCTCAGGTGGAATCAGAT 1863
; :|||:|||||:
Db 473 ProAspThrMetLysArgTyrSerMetTyrLeuThrProLysGlyValArgThrSer 492
Qy 1864 TGG---CCGCTTACACATGGTCAATTTTC-----TCATTTTGGCG 1902
; :|||:|||||:
Db 493 TyrGlnProSerSerProGlyArgPheThrLysGluThrGlnLysAsnPheAsnAsn 512
Qy 1903 ATTTATGATTCA---GGTGGTACTTAGAGAAAGTGTATCAACTGCTAAATCGGTAGAA 1959
; :|||:|||||:
Db 513 LeuTyrGlyThrLysGlyAsnTyrThrSerArgValTrpGluTyrSerSerIleGln 532
; :|||:|||||:
RESULT 13
US-10-525-710-22
; Sequence 22, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zeider, Oskar
; APPLICANT: Kolprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (mety)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
```

```
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-525-710-22
Alignment Scores:
Pred. No.: 0.17 Length: 425
Score: 104.00 Matches: 62
Percent Similarity: 38.8% Conservative: 40
Best Local Similarity: 23.6% Mismatches: 121
Query Match: 1.9% Indels: 40
DB: 6 Gaps: 8
US-10-681-086-1 (1-3156) x US-10-525-710-22 (1-425)
Qy 199 GATCGAGCCCAACAGCATTGCGACAAATTTGGGACAGGAAGCAGCGTTCAGTTTAACG 258
; :|||:|||||:
Db 34 AspAspThrGlnHisGlyAlaAspLeuPheAspLeuLysValProGlyAsnIleTyrThr 53
Qy 259 ACAGGCAATTCGGTCTGGCATGAAAAGCTAGAAAAGAGATGCGCAGCTTTAAATGACA 318
; :|||:|||||:
Db 54 ArgIleMetAsnProThrAsnAspValLeuGluGlnArgValAlaLeuGluGlyGly 73
Qy 319 GAAGCGCGCTGCTGTTTTCGAGCGGTTACTTGGCCCAATGTCGGTGTCTTCATCCCTG 378
; :|||:|||||:
Db 74 ValGlyAlaLeuAlaValAlaSerGlyMetAlaIleThrTyrAlaIleGlnThrVal 93
Qy 379 CCAGAAAAGGAAGATGTCTTTAAAGTGACCGCTCAATCATGCAAGTATGATCGACGGC 438
; :|||:|||||:
Db 94 AlaGluAlaGlyAspAsnIleValSer-----ValAlaLysLeuTyrGlyGly 109
Qy 439 TGGCAGCTTTCTAAGGCTGATACAGTT-----GTTTATCGG 474
; :|||:|||||:
Db 110 ThrTyrAsnLeuLeuAlaHisThrLeuProArgIleGlyIleGlnAlaArgPheAlaLeu 129
Qy 475 CATATTGATATGATGATCTTTGAAAACAAAGCTGAATGAAACACACACGCTTATCAGCGCGT 534
; :|||:|||||:
Db 130 HisAspAspValAlaLeuLeuGluAlaLeuIleAspGlu-----ArgThrLys 145
Qy 535 TTTATCGTAACAGACGAGGATATTTCAGCATGGATGGCAACAATCGCCCTCTTTGATCAGATC 594
; :|||:|||||:
Db 146 AlaValPheCysGluThrIleGlyAsnProAlaGlyAsnIleIleAspLeuGlnAlaLeu 165
Qy 595 ATCTCATTGCGGAACCGCTATCATGCTTCGTGCTGTTGATGATGCCACGACACAGGA 654
; :|||:|||||:
Db 166 AlaAspAlaAlaHisArgHisGlyValProLeuIleValAspAsnThrValAlaThrPro 185
Qy 655 GTTTTGGCGGATTTCGGGACCAAGGACGAGTGAATCTTTGTGTTGTGCGGACATGTTT 714
; :|||:|||||:
Db 186 ValLeu-----CysArgProPheGluHisGlyAlaAspIleVal 198
Qy 715 ATCGGCACCTTAAGCAAAAGCTGTTGGCGCGGAA-----GGAGGTTTTTCGGCA 762
; :|||:|||||:
Db 199 ValHisSerLeuThrLysTyrMetGlyHisGlyThrSerIleGlyGlyIleValVal 218
Qy 763 GGATCAGCGGCTCTTCATCGACTTTTTCGTAACCATGCCAGAACATTTATCTTCAAAACC 822
; :|||:|||||:
Db 219 AspSerGlyLysPhe---AspTrpAlaAlaAsnLysSerArgPheProLeuAsnThr 237
Qy 823 GCTATTCCGCCAGCCAGCTGTGGCGTCTCACGAGGCTTTCACACATCATTTGAAGCCAGC 882
; :|||:|||||:
Db 238 ProAspProSerTyrHisGlyValThrTyrThrGluAlaPheGlyProAla----- 254
Qy 883 AGGGAAGAACGACAGCTTTTATTTCTTATATCATGATGATCAGAGTGAAG 939
; :|||:|||||:
```

```

255 -----AlaPheIleGlyArgCysArgValValProLeuArg 266
QY 940 AATATGGGT 948
Db |||||
267 AsnMetGly 269
Db |||||

RESULT 14
US-10-995-561-776
; Sequence 776, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 776
; LENGTH: 5935
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-776

Alignment Scores:
Pred. No.: 0.556 Length: 5935
Score: 103.50 Matches: 96
Percent Similarity: 36.8% Conservative: 75
Best Local Similarity: 20.4% Mismatches: 127
Query Match: 1.9% Indels: 167
DB: 6 Gaps: 21

US-10-681-086-1 (1-3156) x US-10-995-561-776 (1-5935)
QY 1912 TCAGGTGGACTACTAGAGAAGTGTATCAAACTGCTAAATGGTGAGAGCCCAACCGTTC 1971
Db :::::
1973 AsnGlyGlyValAsnProGluProPheArgAlaThrGlnAsnVal-----Phe 1798
QY 1972 CACGATCGGATTTGCGCTTATCGTAGAGAGCTGTTGATATGACGAGCAATGCGGT 2031
Db :::::
1799 ThrArgGlnLeuCysLeuGluHisAspGluLysLeuValSerTyrLeuSerLeuLeuArg 1818
QY 2032 AATATTGCTGTCGACGACGACCAACATTTCTACCATCTTGACTGTACAGTAGCAATG 2091
Db |||||
1819 AsnIleGluMetArg-----ThrLysGlnIleGlnPro 1829
QY 2092 GCAGGTGCCATGTTGATTGGTCTGCATCATCGCATCTGT----- 2130
Db |||||
1830 LeuGluLeuAsnLeuAlaGluLeuGlnAspLeuLeuCysGlnAlaLysValLeuGluArg 1849
QY 2131 -----TATACAGCAGCGCTTCGCTTAACTGAAGCAGTAAAGCAATCAGAT 2178
Db |||||
1850 GluLeuLysAspLeuThrThrLeuValSerGlnGluLeuGluCysValAsnGlnIle 1869
QY 2179 CTTCTCTTCAGTTATGACCATCTGTCGAGTTCGTA-----ATGTCGTGTCACCTTCC 2232
Db :::::
1870 IleSerGlnProGlnGluValProAlaGlnLeuLeuLysAlaLeuGluLysAspAlaLys 1889
QY 2233 GACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACA 2292
Db :::::
1890 AsnLeuGlnLysSerLeuSerValSerAspThrIrpAsn----- 1903
QY 2293 GAACGACACGGATATATAGTGAGTGTGCAAAACGCATACCATTTTGAAATTCGAAAGCGC 2352
Db |||||
1904 -----SerArgLeuLeu-HisPheGln-AsnAlaValG 1914
QY 2353 CGATTGACTTACCGGATGTGATAGGAACGCTGCTTGAAGCGCGGAGCGGCGC 2412
Db |||||
1914 IuIleGluLysThrLysValLeuAsnGlnHisThrGlnLeuGluGly-----ArgLeuG 1932
QY 2413 ACGATCTGAAT-----ATCGGCACATATTGTATCGTGCAAGCGGACGAGGTCCGCTA 2466
```

```

||||| ::::: ::::: ::::: :::::
1932 InAspLeuArgAlaTrpValGlyAsnLysAsnLeuIleLeuAsnSerLysGly---SerA 1951
QY ACAGACAAGTGGATCAGGTCTAGATGCGGTT-----CAGGAAA 2505
Db ||||| :::::
1951 snSerGluIleAsp-----ValAspSerLeuAsnLeuCysLeuGlnGlnTyrGluAspL 1969
QY TTAAGACAGCAGTATGCAAGATTTGTGCATGCTCTTGCACTG----- 2549
Db :::::
1969 euLysGlnProMetAlaGluArgLysAlaGlnLeuAspAlaLeuAlaPheAspIleGlnP 1989
QY -----TTGAAGCCAGACAGCGGCGAAGCGCTCAAGATG 2583
Db ||||| :::::
1989 hePheIleSerGluHisAlaGlnAspLeuSerProGlnGlnAsnArgGlnMetLeu--- 2007
QY CAGGAGTAGACCGCTATATCAATAATTGTAATACGTCACAGAGAAACCATTTCAAACATCA 2643
Db |||||
2008 -----ArgLeuLeuAsnGlnLeuGlnArgSerPheGlnAspIleL 2021
QY CAACTCA----- 2651
Db
2021 euGluGlnThrAlaAlaGlnValAspAlaLeuGlnGlyHisLeuGlnGlnMetGluGlnG 2041
QY -----CATACATAGATGACAGAGTCAATACGTTGAATCGCAAAAGAA----- 2696
Db |||||
2041 luAlaLeuValLysThrLeuGlnLysGlnGlnAsnThrCysHisGlnGlnLeuGluAspL 2061
QY -----TCGGGGCTGCTCCGTTGAGGCCCATTTATCGGGATG-----AAGGAGACGA 2745
Db |||||
2061 euCysSerTrpValGlyGlnAlaGluArgAlaLeuAlaGlyHisGlnGlyArgThrThrG 2081
QY AACAGAGTGTCAATTGACATCGCCAAAGC-----TTGAAGGCTCTT-----GACG 2790
Db :::::
2081 InGlnAspLeuSerAlaLeuGlnLysAsnGlnSerAspLeuLysAspLeuGlnAspAspI 2101
QY CGGATTCCTCTGTAATTT-----TTGATGCAATGATGGCAGCGCGTTAG 2841
Db :::::
2101 leGlnAsnArgAlaThrSerPheAlaThrValLysAspIleGluGlyPheMetGluG 2121
QY AAGGCGTCAACGAATTAACCCGCTGTATTGTTTAAAGTGTGGCGCTGCTTCGTTTAA 2901
Db :::::
2121 luAsnGlnThrLysLeuSerPro-----ArgGluLeuThrAlaLeuArgGluLysL 2138
QY TCAATCCATCAAAAGAAATTCGCATTTTCGAGAGAGAGAGGTCATTCGCGACATGCG 2961
Db :::::
2138 euHisGlnAlaLysGlu-----GlnTyr-GluAlaLeu 2149
QY AGCCATTAGGCTTTAGCGCGCAAACTCCATTTTTCGCGAGACTACTTAACAACGCGC 3021
Db |||||
2149 InGluGluThrArgValAlaGlnLysGluLeu-----GluGluAlaValThrSerAlaL 2167
QY GGCACAGAGGAGACGAGGATCATAAATGCTGAGTGATTTAGGCTTTGAAGTTCAATCAG 3081
Db |||||
3082 TCGAAGAAATG 3092
QY :::::
2187 leAspAlaLeu 2190
Db :::::

RESULT 15
US-11-055-822-662
; Sequence 662, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
```

```
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 662
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-662

Alignment Scores:
Pred. No.: 0.239 Length: 348
Score: 102.00 Matches: 52
Percent Similarity: 26.4% Conservative: 34
Best Local Similarity: 22.0% Mismatches: 77
Query Match: 1.8% Indels: 73
DB: 7 Gaps: 11

US-10-681-086-1 (1-3156) x US-11-055-822-662 (1-348)
QY 2195 ACCATCTGTCGAGTTCGTATGCTGTCACTTCGACTCTGAGAACTTCGGAAAT 2254
Db 53 ThrValCysGlnGluAlaGlyCys-----ProAsnIleHisGluCysTrpGlu 68
QY 2255 GCCTAGAGAAATTCGGAAATCGGATTACAGAGTGGAGAACGACACCGATATATAGTGG 2314
Db 69 -SerArgGluAlaThrPheLeuIleGlyGlyAlaAsnCysSerArgArg----- 84
QY 2315 ATGTGTCAAAACGCATACCATTTTGAATTCGAAAAGCGCG----- 2354
Db 85 -Cys-----AapPheCysMetIleAsnSerAlaArgProGluProLeuAspArgGlyG1 102
QY 2355 -----ATTGAGTCTTACCG-----ATGGTGAATAAGGAAACGCTGCTT-- 2393
Db 102 uProLeuArgValAlaGluSerValArgGluMetGlnLeuAsnTyrSerThrIleThrG1 122
QY 2394 -----GAAGGCGCG-----AAGCG 2407
Db 122 yValThrArgAspAspLeuAspAspGluGlyAlaTrpLeuTyrSerGluValValArgLy 142
QY 2408 GCGCAGCATCTGAAT-----ATCGGCACATATTGTATCGTGGCAAGCGCAGAGTCC 2461
Db 142 sIleHisGluLeuAsnProHisThrGlyValGluAsnLeuValProAspPheSerGlyLy 162
QY 2462 GTCTAACAGAGAAGTGGATCGGTAGTCGGTTCCGGAATTAAGAGAGCTATGG 2521
Db 162 sLys-----AspLeuLeuGlnGluValPheGluSer----- 172
QY 2522 ACTGAAGATTGTGCATGCTTTGGACTGTTGAAGCCAGACGACGCGAAGCGGCTCAAAGA 2581
Db 173 -----ArgProGluVal----- 176
QY 2582 TGCAGGAGTAGCCGCTAATCATTAATTTGAAATCGTCAACAGAAACCATTCAAACAT 2641
Db 2582 TGCAGGAGTAGCCGCTAATCATTAATTTGAAATCGTCAACAGAAACCATTCAAACAT 2641
```

```
Db 177 -----PheAlaHisAsnValGluThrValProArgIlePheLysArgI1 191
QY 2642 CACAACCTCACATACATACGATGACAGATCAATACGGTTCAAAATCGCAAAAGATCCGG 2701
Db 191 eArgProAlaPheArgTyrGluArgSerLeuAspValIleArgGlnAlaArgAspPheG1 211
QY 2702 GCTGTCTCCGTGTTTCAGGCGCCATTATCGGGATCAAGAGACGAAACACAGGATGTCATTGA 2761
Db 211 yLeuValThrLysSerAsnLeuIleLeuGlyMetGlyGluThrLysGluGluIleThrG1 231
QY 2762 CATGCCCAAAAGCTTGAAGGCTTTGACGCGGATTCATTCCCTCTGTG 2807
Db 231 uAlaLeuGlnAspLeuHisAspAlaGlyCysAspIleIleThrIle 246
```

Search completed: February 10, 2006, 04:07:17
Job time : 57 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2006, 03:00:37 ; Search time 271.6 Seconds
(without alignments)
3211.672 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggttacgagc.....atcggtatgatgcgaattc 3156

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7861189 seqs, 1381955077 residues

Total number of hits satisfying chosen parameters: 15722378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abes/ABSSWEB_spool/US10681086/runat_09022006_140522_11866/app_query.fasta_1
-DB=Pending Patents AA_Main -QFWT=faskan -SUFFIX=n2p.rapm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abses02p
-USER=US10681086 @CGN 1 1 1104 @runat_09022006_140522_11866 -NCPU=6 -ICPU=3
-NO_MMAPP -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG -DRV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents AA_Main:
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US066_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US074_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US075_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US076_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US077_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US078_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US079_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*

28: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
38: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
39: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
40: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
41: /cgn2_6/ptodata/1/paa/US111_COMB.pep.*
42: /cgn2_6/ptodata/1/paa/US112_COMB.pep.*
43: /cgn2_6/ptodata/1/paa/US114_COMB.pep.*
44: /cgn2_6/ptodata/1/paa/US600_COMB.pep.*
45: /cgn2_6/ptodata/1/paa/US601_COMB.pep.*
46: /cgn2_6/ptodata/1/paa/US602_COMB.pep.*
47: /cgn2_6/ptodata/1/paa/US603_COMB.pep.*
48: /cgn2_6/ptodata/1/paa/US604_COMB.pep.*
49: /cgn2_6/ptodata/1/paa/US605_COMB.pep.*
50: /cgn2_6/ptodata/1/paa/US606_COMB.pep.*
51: /cgn2_6/ptodata/1/paa/US607_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1836	33.3	389	27	US-09-791-537-52950
2	1836	33.3	389	33	US-10-369-493-23255
3	1836	33.3	389	47	US-60-360-039-23255
4	1836	33.3	389	50	US-60-638-099-24877
5	1676	30.4	371	50	US-60-638-099-43392
6	1463	26.5	279	18	US-08-827-356-3622
7	1463	26.5	279	26	US-09-611-529-6706
8	1463	26.5	279	29	US-09-950-084-6706
9	1463	26.5	279	39	US-10-914-020-6706
10	1350	24.5	256	27	US-09-791-537-126128
11	1344	24.4	256	27	US-09-791-537-80888
12	1333	24.2	253	1	PCT-US97-07698-2
13	1333	24.2	253	1	PCT-US97-07853-4
14	1333	24.2	253	16	US-08-642-045A-4
15	1333	24.2	253	27	US-09-791-537-104530
16	1333	24.2	253	33	US-10-320-800-75
17	1333	24.2	253	37	US-10-718-628-11
18	1327	24.1	253	27	US-09-791-537-99729
19	1327	24.1	253	27	US-09-791-537-104529
20	1324	24.0	253	27	US-09-791-537-128869
21	1318	23.9	253	27	US-09-791-537-21870
22	1318	23.9	253	37	US-10-718-628-1
23	1286.5	23.3	379	1	PCT-US04-37204-4913
24	1282	23.2	253	37	US-10-718-628-12
25	1277	23.1	253	37	US-10-718-628-14
26	1268	23.0	253	37	US-10-718-628-15
27	1267	23.0	253	37	US-10-718-628-2
28	1267	23.0	253	37	US-10-718-628-13
29	1265	22.9	253	37	US-10-718-628-20
30	1262	22.9	253	37	US-10-718-628-17
31	1262	22.9	253	37	US-10-718-628-18
32	1260	22.8	253	37	US-10-718-628-16
33	1256	22.8	253	37	US-10-718-628-19
34	1222	22.1	253	37	US-10-718-628-3
35	1222	22.1	333	1	PCT-US04-37204-4915
36	942	17.1	387	19	US-08-935-263-6
37	942	17.1	387	50	US-60-638-099-42172
38	929	16.8	389	27	US-09-791-537-12753
39	859.5	15.6	388	50	US-60-638-099-26324

40 851 15.4 338 19 US-08-935-263-8 Sequence 8, Appli
41 843 15.3 392 1 PCT-US03-09585-26 Sequence 26, Appl
42 843 15.3 392 27 US-03-731-537-44892 Sequence 44892, A
43 843 15.3 392 33 US-10-369-493-23135 Sequence 23135, A
44 843 15.3 392 35 US-10-507-720-26 Sequence 26, Appl
45 843 15.3 392 47 US-60-360-039-23135 Sequence 23135, A

ALIGNMENTS

RESULT 1
US-09-791-537-52950
; Sequence 52950, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonmix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52950
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-791-537-52950

Alignment Scores:
Pred. No.: 7,45e-168 Length: 389
Score: 1836.00 Matches: 359
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 33.3% Indels: 0
DB: 27 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-791-537-52950 (1-389)

QY 22 TTGAAGATTGATCTCTGTTAAACAGCGGTTAGACAGAAATGAAAGAACCGCGGTACAT 81
:::
Db 1 MetLysleAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20

QY 82 CTAACCTCGGTCATGATCGAGCGCGGTCAGAGAGCAATATTGATGCGCAAAAT 141
Db 21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40

QY 142 CAAACGGTCTGCTCTCAACAAATATTATTAGGCTCGCAAGCATAGACGTTTGATCGAT 201
Db 41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgLeuIleAsp 60

QY 202 CGAGCCCAACAGATTGACGAATTTGGACAGGACGCGGTTACGTTTAAACGACA 261
Db 61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80

QY 262 GCAATTCGGTCTGCGATGAAAGAGTTCAGAAAGAGATTCGCGCTTAAACTGCAGAA 321
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100

QY 322 GGGCGCTCTGCTTTTCAGCGGTTACTTTGGCAATGTCGGGTGCTCTTCATCTTCGCA 381
Db 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120

QY 382 GAAAGGAGATGTCTATTTAAGTACCAAGCTCAATCATGCAAGTATGATCGACGCTGC 441
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140

QY 442 CGACTTCTTAGGCTGTACAGTCTTTATCGGCATATTGATGAATGATCTTCAAAAC 501
Db 141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160

QY 502 AAGCTGAATGAAACACAGCGTTTATCAGCGCGCTTTTATCGTAAACAGCGGAGTATTACG 561

Db 161 LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
QY 562 ATGATGACGACAATCGCCCTCTTGTATGATCAGATCATCTCCTTGGGAAACGCTATCATGCC 621
Db 181 MetAspGlyThrIleAlaProLeuAspGlnIleIleSerLeuAlaLysArgTyrHisAla 200

QY 622 TTCGTGTCGTTGATGATGCCACGACGAGAGTTTGGCGGATTCGGGCAACAGGAACG 681
Db 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220

QY 682 AGTGAATACTTTGCTGTTTGTCCGACATTTTATCGGCACCTTTAAGCAAAAGCTTTGGC 741
Db 221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240

QY 742 CGGGAAGAGGTTTGGCGAGGATCAGCGGCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260

QY 802 AGAACATTATTCTTTCAAAACGCTATTCCGCGACGCTGTCGGCTGCTCACGAGGCT 861
Db 261 ArgThrPheIlePheGlnThrAlaIleProProIleAspCysAlaAlaHisGluAla 280

QY 862 TTCAACATCATTTGAAGCCAGCAGGAGGAAAAACGACAGCTTTTATTTCTTATATCAGCATG 921
Db 281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300

QY 922 ATCAGAACCATCTGAAGAATATCGGTTATGCTGTGTAAGGAGATCACACACCGATTATT 981
Db 301 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 320

QY 982 CCTGTAGTCATTGCGCATGCCATATAAAACGCTCTATTGCTGAAAAACTGCAGGGCAAG 1041
Db 321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340

QY 1042 GGAATTTATGCTCTGCTCATTCGCGCCGCAACCGTTGCGCGGTTGAAAGCCGGATTCCA 1101
Db 341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360

RESULT 2
US-10-369-493-23255
; Sequence 23255, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23255
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-369-493-23255

Alignment Scores:
Pred. No.: 7,45e-168 Length: 389
Score: 1836.00 Matches: 359
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 33.3% Indels: 0
DB: 27 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-369-493-23255 (1-389)

QY 22 TTGAAGATTGATCTCTGTTAAACAGCGGTTAGACAGAAATGAAAGAACCGCGGTACAT 81


```

:::|||||
1 MetLysileAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20
82 CGTAACCTGCGGTCAATGATGGAGCCGCGTTCAGAGAGGATATTTGATGGCGAAAT 141
21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
142 CAACCGGTCTGCTCTCAAAACAATTTATAGGCGTCGCAAGCGATAGACGTTTCATCGAT 201
41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspAsgArgLeuIleAsp 60
202 GCACCCCAACAGCATTTGCAGCAATTTGGGACAGAGACGCGGTTCACGTTTAAACGACA 261
61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThr 80
262 GGCAATTCGGTCTGCGCATGAAAGCTAGAAAGAGATTGCCAGCTTTAAACTGCACAGAA 321
81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100
322 CGCGCCCTGCTGTTTCGAGCGGTACTTGGCCCAATGTCGGTGTCTTTCATCTTGCCA 381
101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
382 GAAAGGAAGATGCTATTTAAGTGACCGATCAATCATGCAAGTATGATCGACGGCTGC 441
121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
442 CGACTTCTAAGGCTGATACAGTTGTTTATCGGCATATTCGATATGATGATCTTGAAC 501
141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160
502 AAGCTGTAATGAAACACACAGCGTTATCAGCGCGTTCGTAACAGACGAGTATTCAGC 561
161 LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
562 ATGGATGSCACAATCGCCCTCTTGATCAGATCATCTCACTTGGCAACGCTATCATGCC 621
181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTyrHisAla 200
622 TTCGTGCTGCTGTGTTGTTGTCGACATTCGTCACCTTAAAGCAAGCTGTGGC 741
221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
742 CGGAAGAGGAGGTTTGGCGCAGGATCAGCGGTCTTCATCGACTTTTTCGTAACCATGCC 801
241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
802 AGACATTTATCTTTCAACCGCTATTCGCGCCAGCAGCTGTGGCTGCTCAGAGGCT 861
261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaAlaHisGluAla 280
862 TTCACATCATTTGAAGCCAGCAGGGAACACACAGCTTTTATTTCTATATCAGCATG 921
281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
922 ATCAGAACCATGCTCTGAAGAATATGGTTATGTGTGTAAGAGAGATCATCACACCGATATT 981
301 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 320
982 CCTGTAGTCATGCGGATGCCATAAACGGTCTCTATTTGCTGAAAACCTCAGGCGNAG 1041
321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
1042 GGAATTTATGCTCTGCGCATTCGCGCCGCAACCGTTGCGCGGTGGAAGCCGATTCGA 1101
341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360
```

RESULT 3

US-60-360-039-23255

; Sequence 23255, Application US/60360039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCES: 38-10(52052)A

; CURRENT APPLICATION NUMBER: US/60/360,039

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 23255

; LENGTH: 389

; TYPE: PRT

; ORGANISM: Bacillus subtilis

US-60-360-039-23255

Alignment Scores:

Pred. No.: 7,45e-168 Length: 389

Score: 1836.00 Matches: 359

Percent Similarity: 100.0% Conservative: 1

Best Local Similarity: 99.7% Mismatches: 0

Query Match: 33.3% Indels: 0

DB: 47 Gaps: 0

US-10-681-086-1 (1-3156) x US-60-360-039-23255 (1-389)

```

Qy 22 TTGAAGATTGATCTCTCGTTTAAACGACGCGTTAGACAGATGAAAGAGCGCGGTACAT 81
Db 1 MetLysileAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20
Qy 82 CGTAACCTGCGGTCAATGATGGAGCCGCGTTCAGAGAGGATATTTGATGGCGAAAT 141
Db 21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
Qy 142 CAACCGGTCTGCTCTCAAAACAATTTATAGGCGTCGCAAGCGATAGACGTTTCATCGAT 201
Db 41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspAsgArgLeuIleAsp 60
Qy 202 GCACCCCAACAGCATTTGCAGCAATTTGGGACAGAGACGCGGTTCACGTTTAAACGACA 261
Db 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThr 80
Qy 262 GGCAATTCGGTCTGCGCATGAAAGCTAGAAAGAGATTGCCAGCTTTAAACTGCACAGAA 321
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100
Qy 322 CGCGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTCGGTGTCTTTCATCTTGCCA 381
Db 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
Qy 382 GAAAGGAAGATGCTATTTAAGTGACCGATCAATCATGCAAGTATGATCGACGGCTGC 441
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
Qy 442 CGACTTCTAAGGCTGATACAGTTGTTTATCGGCATATTCGATATGATGATCTTGAAC 501
Db 141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160
Qy 502 AAGCTGTAATGAAACACACAGCGTTATCAGCGCGTTCGTAACAGACGAGTATTCAGC 561
Db 161 LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
Qy 562 ATGGATGSCACAATCGCCCTCTTGATCAGATCATCTCACTTGGCAACGCTATCATGCC 621
Db 181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTyrHisAla 200
Qy 622 TTCGTGCTGCTGTGTTGTTGTCGACATTCGTCACCTTAAAGCAAGCTGTGGC 681
Db 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
```

```
QY 682 AGTGAATACATTTTGGTGTGTTTCCGACATTTGTTATCGGCACCTTAAGCAAAAGCTGTGGC 741
Db |||||
QY 221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
Db |||||
QY 742 GCGAAGAGAGTTTGGCGGAGATCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db |||||
QY 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAenHisAla 260
QY 802 AGACATTTACTTTTCAAAACCGCTATTCCGCCAGCAGCTGTGGCGTCTCAGAGGCT 861
Db |||||
QY 261 ArgThrPheIlePheGlnThrAlaIleProAlaSerCysAlaAlaAlaHisGluAla 280
QY 862 TTCACATCATTTGAAGCCAGCAGGCAAAACACACAGCTTTTATTTCTTATATCAGCATG 921
Db |||||
QY 281 PheAenIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
QY 922 ATCAGAACAGCTCTGAAGAAATATGGGTATGTGTGTAAGAGATCATCACACCGCATTTAT 981
Db |||||
QY 301 IleArgThrSerLeuLysAenMetGlyTyrValValLysGlyAspHisThrProIleIle 320
QY 982 CCTGTAGTCATTGGCGATGCCATAAACGGTCTCTATTGCTGAAAACTCGAGGCAAG 1041
Db |||||
QY 321 ProValIleIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
QY 1042 GGAATTTATGCTCTGCCATTCCGCCGCCAACCGTTGCGCGGTGAAAGCGCATTCGA 1101
Db |||||
QY 341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360
Db |||||
RESULT 4
US-60-638-099-24877
; Sequence 24877, Application US/60638099
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: Transgenic Plants with Enhanced Agronomic Phenotypes
; FILE REFERENCE: 38-21(53720)
; CURRENT APPLICATION NUMBER: US/60/638,099
; CURRENT FILING DATE: 2004-12-21
; NUMBER OF SEQ ID NOS: 48056
; SEQ ID NO 24877
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-60-638-099-24877
```

```
Alignment Scores:
Pred. No.: 7,45e-168 Length: 389
Score: 1836.00 Matches: 359
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 33.3% Indels: 0
DB: 50 Gaps: 0
```

US-10-681-086-1 (1-3156) x US-60-638-099-24877 (1-389)

```
QY 22 TTGAAGATTGATTCCTGGTTAAACGAGCGGTAGACAGAATGAAGAAGCCGCGTACAT 81
Db |||||
QY 1 MetLysIleAspSerTrpLeuAenGluArgLeuAspArgMetLysGluAlaGlyValHis 20
QY 82 CTTAACCTCGCGTCAATGATGAGCGCGGTTCAGAGAGAAATATTGATGGCGAAAT 141
Db |||||
QY 21 ArgAenLeuArgSerMetAspGlyAlaProValProGluArgAenIleAspGlyGluAen 40
QY 142 CAAACGGTCTGCTCTCAACAAATTTATAGGCTCGCAGCGATAGACGTTGATGAT 201
Db |||||
QY 41 GlnThrValTrpSerSerAenTyrLeuGlyLeuAlaSerAspArgArgLeuIleAsp 60
QY 202 GCAGCCCAACAGCATTCGACAAATTTGGCAGCAGGAAGCAGCGGTTCACGTTTAAACGACA 261
Db |||||
QY 61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
QY 262 GCAATTCGGTCTGGCATGAAAGCTAGAAAGAGATTGCGAGCTTTAAACTGCAGAA 321
Db |||||
QY 81 GlyAenSerValTrpHisGluLysLeuGluLysLysIleAlaSerPheLysLeuThrGlu 100
Db |||||
```

```
QY 322 GCGGCCCTGCTGTTTTCGAGCGGTTACTTGCCCAATGTGCGTGTCTTTCATCTTCCCA 381
Db |||||
QY 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAenValGlyValLeuSerSerLeuPro 120
QY 382 GAAAAGAGAGATGTCATTTTAAAGTGACAGCTCAATCATGCAAGTATGATCGAGCGTGC 441
Db |||||
QY 121 GluLysGluAenValIleLeuSerAspGlnLeuAenHisAlaSerMetIleAspGlyCys 140
QY 442 CGACTTTCTAAGGCTGATACAGTTGTTTATCGGCATATGATATGATATGATCTTGAAC 501
Db |||||
QY 141 ArgLeuSerLysAlaAspThrValTyrArgHisIleAspMetAenAspLeuGluAen 160
QY 502 AAGCTGAATGAAACACACAGCGTTATCAGCGCGTATTATCGTAAACAGACGAGTATTCAGC 561
Db |||||
QY 161 LysLeuAenGlnThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
QY 562 ATGATGCGCAATCGGCCCTCTTTGATCAGATCATCTCATCTTGGGAAACGCTATCATGCC 621
Db |||||
QY 181 MetAspGlyThrIleAlaProLeuAspGlnIleIleSerLeuAlaLysArgTyrHisAla 200
QY 622 TTCGTGCTGCTGTGATGATGCCACGCAACAGGAGCTTTTGGCGGATTCGGGACAAAGGACG 681
Db |||||
QY 201 PheValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
QY 682 AGTCAATACTTTGCTGTTTGTCCGACATTTGTTATCGGCACCTTAAGCAAAAGCTGTGGC 741
Db |||||
QY 221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
QY 742 GCGAAGAGAGTTTGGCGGAGATCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db |||||
QY 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAenHisAla 260
QY 802 AGACATTTACTTTTCAAAACCGCTATTCCGCCAGCAGCTGTGCGGTCTCAGAGGCT 861
Db |||||
QY 261 ArgThrPheIlePheGlnThrAlaIleProAlaSerCysAlaAlaAlaHisGluAla 280
QY 862 TTCACATCATTTGAAGCCAGCAGGCAAAACACACAGCTTTTATTTCTTATATCAGCATG 921
Db |||||
QY 281 PheAenIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
QY 922 ATCAGAACAGCTCTGAAGAAATATGGGTATGTGTGTAAGAGATCATCACACCGCATTTAT 981
Db |||||
QY 301 IleArgThrSerLeuLysAenMetGlyTyrValValLysGlyAspHisThrProIleIle 320
QY 982 CCTGTAGTCATTGGCGATGCCATAAACCGTCTCTATTGCTGAAAACTCGAGGCAAG 1041
Db |||||
QY 321 ProValIleIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
QY 1042 GGAATTTATGCTCTGCCATTCCGCCGCCAACCGTTGCGCGGTGAAAGCGCATTCGA 1101
Db |||||
QY 341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360
Db |||||
```

RESULT 5

US-60-638-099-43392
; Sequence 43392, Application US/60638099
; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: Transgenic Plants with Enhanced Agronomic Phenotypes
; FILE REFERENCE: 38-21(53720)
; CURRENT APPLICATION NUMBER: US/60/638,099
; CURRENT FILING DATE: 2004-12-21
; NUMBER OF SEQ ID NOS: 48056
; SEQ ID NO 43392
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-60-638-099-43392

```
Alignment Scores:
Pred. No.: 2.36e-152 Length: 371
Score: 1676.00 Matches: 330
Percent Similarity: 93.3% Conservative: 6
```

Best Local Similarity: 91.7% Mismatches: 6
Query Match: 30.4% Indels: 18
DB: 50 Gaps: 1

US-10-681-086-1 (1-3156) x US-60-638-099-43392 (1-371)

QY 22 TTCAAGATTGATCTCTGTTTAAACGAGCGGTAGACAGANTGAAGAAGCGCGGTACAT 81
:::|||||
Db 1 MetLysIleAspSerTrpLeuAsnAspArgLeuAspIleAlaLysGluAlaGlyValHis 20
|||
QY 82 CGTAACTCGGTCAATGATCGAGCGCGGTTCAGAGAGGAATATTGATGCGCAAAAT 141
|||
Db 21 ArgAsnLeuArgSerMetAsnGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
|||
QY 142 CAAACGGTCTGCTCTCAAAACAATATTATTAGGGCTCGCAAGCGATPAGACGTTTGCAT 201
|||
Db 41 GlnThrValTrpSerSerAsnAsnTrpLeuGlyLeuAlaSerAspArgArgLeuIleAsp 60
|||
QY 202 GCAGCCCAACAGCATTCAGCAATTTGGGACAGAGACGCGGTTCAGTTTAAACGACA 261
|||
Db 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThr 80
|||
QY 262 GGCAAATTCGGTCTGGCATCAAAAGCTAGAAAAGAGATTGCCAGCTTTAAACTCACAGAA 321
|||
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysArgThrGlu 100
|||
QY 322 GCGGCCCTCTGCTTTTCGAGCGGTACTTGGCCAAATGTCGGTGTCTCTTTCATCTGCCA 381
|||
Db 101 AlaAlaLeuLeuPheSerSerGlyTrpLeuAlaAsnValGlyValLeuSerSerLeuPro 120
|||
QY 382 GAAAGGAGATGTCATTTAAGTGACACGCTCAATCATGCAAGTATGATCATCGAGGCTGC 441
|||
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerIleIleAspGlyCys 140
|||
QY 442 CGACTTCTTAAGCTCATACAGTCTGTTTATCGCATATTGATGATGAATGATCTTGAAGAC 501
|||
Db 141 ArgLeuSerLysAlaAspThrValValTrpArgHisIleAspMetAsnAspLeuGluAsn 160
|||
QY 502 AAGCTCAATGAACAACAGCGGTATCAGCCCGCTTTTATCGTAACAGACGAGTATTACAGC 561
|||
Db 161 LysLeuAsnGluThrGlnArgTrpGlnArgArgPheIleValThrAspGlyValPheSer 180
|||
QY 562 ATGATGGACAAATCGCCCTCTTGATCAGATCATCTCACTTGCAGAACGCTATCATGCC 621
|||
Db 181 MetAspGlyThrIleValProLeuAspGlnIleIleSerLeuAlaLysArgTrpHisAla 200
|||
QY 622 TTCGTGCTGTGATGATCCCAACGACAGAGTTTGGCGGATTCGGGACAGGACG 681
|||
Db 201 PheValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyArgGlyThr 220
|||
QY 682 AGTGAATATCTTGTGTGCCAGATGTATTCCGCCACTTAAAGCAAGCTGTGTGGC 741
|||
Db 221 SerGluTrpPheGlyValTrpProAspIleValIleGlyThrLeuSerLysAlaValGly 240
|||
QY 742 GCGAAGAGAGTTTTCGCGCAGATCAGCGGTCTTCATCGACTTTTTCGTAACCATGCC 801
|||
Db 241 ThrGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
|||
QY 802 AGAACATTTATCTTTCAAACCGCTATTCGCCAGCAGCTGTGCGGCTCTCACGAGCT 861
|||
Db 261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaHisGluAla 280
|||
QY 862 TTCACATCATTTGAAGCCAGCAGGGGAAAAACGACAGCTTTTATTTCTTATATCAGCATG 921
|||
Db 281 PheAsn----- 282
|||
QY 922 ATCAGAACCATGCTGAAGATATGCGTTATGTTGGTGAAGAGATCACACACCGATTATT 981
|||
Db 283 IleArgThrSerLeuLysAsnMetGlyTrpValValLysGlyAspHisThrProIleIle 302
|||
QY 982 CTTGTAGTATTCGGCATGCCCAATAAAACGGTCTCTATTTCGTGAAAAAATCGACGGCGAAG 1041
|||
Db 303 ProValValIleGlyAspAlaHisLysThrValIlePheAlaGluLysLeuGlnGlyLys 322
|||

Db 121 AspTrpProLeuThrHisGlyGlnPheSerIleLeuProIleTyrAspSerGlyGly 140
QY 1921 TACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAAAAGCTTCCACGATGCG 1980
Db 141 TyrLeuGluLysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAla 160
QY 1981 ATTTGTGCGCTTATCGTAGAAGAGCTGTTTGAATATATGCGGCAAAATGGCGTAATATTCGT 2040
Db 161 IleCysAlaLeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArg 180
QY 2041 GTCACAGGACCGACACATTTCTACCATCTTCCACTGCTCAGGTAGCAATGGCAGGTGCC 2100
Db 181 ValGlnGlyProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAla 200
QY 2101 ATGTTCATGCTGTCATCATCTGTCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
Db 201 MetLeuIleGlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGlu 220
QY 2161 GCAGTTAAGCAATCAGATCTTCTCAGGTTATGACCATCTGTCGACGTTCTGTAATGTCT 2220
Db 221 AlaValLysGlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSer 240
QY 2221 GGTCAACTTCCGACCTCGAAGACTTCTGGAATCGTAGAATTTCTGGAATGGATT 2280
Db 241 GlyGlnLeuSerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIle 260
QY 2281 CAGGAGTGGACAGAAACACACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 261 GlnGluTrpThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 279

RESULT 8

US-09-950-084-6706
; Sequence 6706, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US2
; CURRENT APPLICATION NUMBER: US/09/950,084
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 6706
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-6706

Alignment Scores: 9.22e-132 Length: 279
Pred. No.: 1463.00 Matches: 279
Score:

Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 26.5% Indels: 0
DB: 29 Gaps: 0
US-10-681-086-1 (1-3156) x US-09-950-084-6706 (1-279)
QY 1501 CTGAATAAGAACGGTCTCTCCAAATATTCTTATTTAGAAAAGCAAAATCTAAATATATCT 1560
Db 1 LeuAsnLysAsnGlyAlaLeuGlnIlePheLeuPheArgLysAlaAsnLeuLysLeuSer 20
QY 1561 GAARAGGGATCAGATAGTGAATCGACCAATAATATCTAGACAGAGAAAGAAAGTGAAG 1620
Db 21 GluLysGlyMetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLys 40
QY 1621 ATTGTTTCATGAATAAAGGAAACGAATATTGGATAAATATGGGATCATGTTAAGGCTATT 1680
Db 41 IleValHisGluLysGluArgIleLeuAspLysTyrGlyAspValLysAlaIle 60
QY 1681 GGTGTTTATGGCTCTCTTCGTCTCAGACTGATGAGCCCTTATTCGGATATTGAGATGATG 1740
Db 61 GlyValTyrGlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMet 80
QY 1741 TGTGTCATCTCAACAGAGAACGACGAGTTTCAGCCATGAATGGCAACCCGCTGAGTGAAG 1800
Db 81 CysValMetSerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLys 100
QY 1801 GTGGAAGTCAATTTTGTATAGCGAAGAGATTCTACTAGATTATTCATCTCAGGTGGAAATCA 1860
Db 101 ValGluValAsnPheAspSerGluLysLeuLeuAspTyrLysSerGlnValGluSer 120
QY 1861 GATTGCGCCCTTACACATGTCATGTTTCTCTATTTTCCGATTTATGATTCAGGTGGA 1920
Db 121 AspTrpProLeuThrHisGlyGlnPheSerIleLeuProIleTyrAspSerGlyGly 140
QY 1921 TACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAAAAGCTTCCACGATGCG 1980
Db 141 TyrLeuGluLysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAla 160
QY 1981 ATTTGTGCGCTTATCGTAGAAGAGCTGTTTGAATATGACGCAAAATGGCGTAATATTCGT 2040
Db 161 IleCysAlaLeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArg 180
QY 2041 GTGCAAGGACCGACACATTTCTACCATCTTCCACTGCTGACGTAGCAATGGCAGGTGCC 2100
Db 181 ValGlnGlyProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAla 200
QY 2101 ATGTTGATTTGCTGTCATCATCTGTCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
Db 201 MetLeuIleGlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGlu 220
QY 2161 GCAGTTTAAAGCAATCAGATCTTCTTCAGGTTATGACCATCTGTCGAGTTCTGTAATGTCT 2220
Db 221 AlaValLysGlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSer 240
QY 2221 GGTCAACTTCCGACCTCGAAGAACTTCTGGAATCGTAGAATTTCTGGAATGGATT 2280
Db 241 GlyGlnLeuSerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIle 260
QY 2281 CAGGAGTGGACAGAACGACGAGTATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 261 GlnGluTrpThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 279
RESULT 9
US-10-914-020-6706
; Sequence 6706, Application US/10914020
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr. et al.
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US1
; CURRENT APPLICATION NUMBER: US/10/914,020
; PRIOR FILING DATE: 2004-08-06
; PRIOR APPLICATION NUMBER: US 09/417,811

```
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 6706
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-914-020-6706

Alignment Scores:
Pred. No.: 9,22e-132 Length: 279
Score: 1463.00 Matches: 279
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 26.5% Indels: 0
DB: 39 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-914-020-6706 (1-279)

QY 1501 CTGAATAGACGGTCTCCCAATATCTTATTAGAAAGCAATCTAATATTCT 1560
Db 1 LeuAnLysAeNGlyAlaLeuGlnIlePheLeuPheArgLysAlaAenLeuLysLeuSer 20

QY 1561 GAAAGGGAATGAGATAGTGAATGAGCAATATATGACTAGAGAAAGAAATGAAG 1620
Db 21 GluLysGlyMetArgIleValAenGlyProIleIleMetThrArgGluGluArgMetLys 40

QY 1621 ATTGTTCAATAATTAAGGAACGAATATTGATAAATATGGGATGATGTTAAGCTATT 1680
Db 41 IleValHisGluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIle 60

QY 1681 GGCTTTTATGCTCTCTGGTCTGACACTGATGGCCCTATTTCGGATATTGAGATGATG 1740
Db 61 GlyValTyrGlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMet 80

QY 1741 TGTGTCATGTCAACAGAGAAAGCAGAGTTAGCCATGAATGACAAACCGGTGAGTGAAG 1800
Db 81 CysValMetSerThrGluGluAlaGluPheSerHisGluIleThrThrGlyGluIleTrpLys 100

QY 1801 GTGAAGTGAATTTGATAGCAAGAGATTTCTACTAGATTATGCAATCTCAGGTGGAATCA 1860
Db 101 ValGluValAenPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSer 120

QY 1861 GATTGGCCGCTCACATGGTCAATTTTCTATTTCGATTTTCCGATTTATGATTTCAGGTGGA 1920
Db 121 AspTrpProLeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGly 140

QY 1921 TACTTAGAAGTGTATCAAACTGCATAATCGGTAGAACCCAAACGTTCCAGATGCG 1980
Db 141 TyrLeuGluLysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAla 160

QY 1981 ATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATATCGAGCAAAATCGCGTAATATTCGT 2040
Db 161 IleCysAlaLeuIleValGluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArg 180
```

```
QY 2041 GTCAAGGACCGCAACACATTTTACCATCTCTGACTGTACAGGTAGCAATGCGAGTGCC 2100
Db ValGlnGlyProThrThrPheLeuProSerLeuThrValGlnValAlaMeAlaGlyAla 200

QY 2101 ATGTTGATTGGTCGCATCATCGCATCTGTTATACGACGACGCGCTTCGGTCTTAACGTAA 2160
Db MetLeuIleGlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGlu 220

QY 2161 GCAGTTAAGCAATCAGATCTTCTTCAGGTATGACCATCTGTGCCAGTTCGTAAATGCTCT 2220
Db AlaValLysGlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSer 240

QY 2221 GGTCAACTTTCCGACTCTGAGAACTTCTGAAATCGCTAGAGAATTTCTGGAATGGGATT 2280
Db GlyGlnLeuSerAspSerGluLysLeuLeuGluSerLeuGluAenPheTrpAsnGlyIle 260

QY 2281 CAGGAGTGGACAGAACGACACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db GlnGluTrpThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 279

RESULT 10
US-09-791-537-126128
; Sequence 126128, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 126128
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-791-537-126128

Alignment Scores:
Pred. No.: 7.98e-121 Length: 256
Score: 1350.00 Matches: 256
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 24.5% Indels: 0
DB: 27 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-791-537-126128 (1-256)

QY 1570 ATGAGATAGTGAATGAGCAATTAATGACTAGAGAAAGAAATGAAGATTGTTTCAT 1629
Db 1 MetArgIleValAenGlyProIleIleMetThrArgGluGluArgMetLysIleValHis 20

QY 1630 GAAATTAAGCAACGAATATTGGATAATATGGGGATGATGTTAAGCGTATTGGTGTATT 1689
Db 21 GluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyr 40

QY 1690 GGCTCTCTTGGTCTGCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTCTATG 1749
Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMet 60

QY 1750 TCACAGAGAAAGCAGAGTTTCAGCCATGAATGGAACACCGGTGAGTGGAGGTGGAAGTG 1809
Db 61 SerThrGluGluAlaGluPheSerHisGluIleThrThrGlyGluIleValGluVal 80

QY 1810 AATTTTTCATAGCAAGAGATTTCTACTAGATTATGATTCCTCAGGTGGATCAGATTGGCCG 1869
Db 81 AsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100

QY 1870 CTTACACATCGTCAATTTTTTCTCTATTTCGCCATTTTATGATTTCAGGTGGATCATTAGAG 1929
Db 101 LeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120
```


QY 1930 AAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCGATTGTGCC 1989
DB 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140
QY 1990 CTATCTAGAGAGCTGTTGAATATGACAGGCAAAATGGCGTAAATATCTGTGCAAGGA 2049
DB 141 LeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGly 160
QY 2050 CCACCAACATTTTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCGCATGTTGATT 2109
DB 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180
QY 2110 GGTCTGCATCATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACGAGCAAGTTAAG 2169
DB 181 GlyLeuHisArgIleCysTyrThrSerAlaSerValLeuThrGluAlaValLys 200
QY 2170 CAATCATGATCTCTCAGGTATGACCACTGTGCGCAGTTCGTAATGCTGTGCAACTT 2229
DB 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220
QY 2230 TCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGG 2289
DB 221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240
QY 2290 ACAGACGACACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
DB 241 ThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 256

RESULT 11
US-09-791-537-80888
; Sequence 80888, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 80888
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Bacillus sp
US-09-791-537-80888

Alignment Scores:
Pred. No.: 3,04e-120 Length: 256
Score: 1344.00 Matches: 255
Percent Similarity: 99.6% Conservatives: 0
Best Local Similarity: 99.6% Mismatches: 1
Query Match: 24.4% Indels: 0
DB: 27 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-791-537-80888 (1-256)

QY 1570 ATGAGATAGTGAATGACCAATAATATGACTAGAGAGAGAAATGAAGATTGTTTCAT 1629
DB 1 MetArgIleValAsnGlyProIleLeuMetThrArgGluGluArgMetLysIleValHis 20
QY 1630 GAAATTAAGNACGAATATGGAATAATATGGGATCATGTTAAGGCTATTGGTGTAT 1689
DB 21 GluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyr 40
QY 1690 GGTCTCTTGTCTCAGACTGATGGGCCCTATTTCGGATATTTCAGATCATGTGTGTCATG 1749
DB 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMet 60
QY 1750 TCAACAGAGGAACGAGATTCCAGCCATGAATGGCAACCGGTGAGTGGAAAGGTGGAAAGTG 1809

DB 61 SerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluVal 80
QY 1810 AATTTGATAGCCAGAGAGATTCTACTAGATTATGTCATCTCAGGTGGAAATCAGATTGCGCG 1869
DB 81 AspPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100
QY 1870 CTATCATGCTCAATTTTCTCTATTTTCCGATTTATGATTCAGGTGATATCTAGAG 1929
DB 101 LeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120
QY 1930 AAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCGATTGTGCC 1989
DB 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140
QY 1990 CTATCTAGAGAGCTGTTGAATATGACAGGCAAAATGGCGTAAATATCTGTGCAAGGA 2049
DB 141 LeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGly 160
QY 2050 CCACCAACATTTTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCGCATGTTGATT 2109
DB 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180
QY 2110 GGTCTGCATCATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACGAGCAAGTTAAG 2169
DB 181 GlyLeuHisArgIleCysTyrThrSerAlaSerValLeuThrGluAlaValLys 200
QY 2170 CAATCATGATCTCTCAGGTATGACCACTGTGCGCAGTTCGTAATGCTGTGCAACTT 2229
DB 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220
QY 2230 TCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGG 2289
DB 221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240
QY 2290 ACAGACGACACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
DB 241 ThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 256

RESULT 12
PCT-US97-07698-2
; Sequence 2, Application PC/TUS9707698
; GENERAL INFORMATION:
; APPLICANT: Pachuk, Catherine J.
; APPLICANT: Samuel, Mano
; APPLICANT: Zurawski, John A.
; APPLICANT: Satishchandran, C.
; TITLE OF INVENTION: CHAIN REACTION CLONING
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/07698
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/642,045
; FILING DATE: 06-MAY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: APOL-0294
; TELECOMMUNICATION INFORMATION:


```

? TELEPHONE: 215-568-3100
?
? TELEFAX: 215-568-3439
?
? INFORMATION FOR SEQ ID NO: 1
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 253 amino acids
?
? TYPE: amino acid
?
? TOPOLOGY: linear
?
? MOLECULE TYPE: protein
?

```

Alignment Scores:		
Pred. No.:	3,536-119	253
Score:	1333.00	252
Percent Similarity:	100.0%	Matches: 1
Best Local Similarity:	99.6%	Mismatches: 0
Query Match:	24.2%	Indels: 0
DB:	1	Gaps: 0

US-10-681-086-1 (1-3156) x PCT-US97-07698-2 (1-253)

1579	Qy	GTGNAATGGACCAATAATAATGCTAGAGAGAGAAAGATGAAGATTGTTTCATGAAATTAAG	1638
:			
:			
1	Db	MetAsnGlyProIleIleMetThrArgGluGluArgMetIlysIleValHisGluIleLys	20
:			
1639	Qy	GAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTTTATGGCTCCTCT	1698
:			
:			
21	Db	GluArgIleuAspLysIlyrGlyAspAspValIysAlaIleGlyValIyGlySerLeu	40
:			
:			
1699	Qy	GGTCGTGAGACTGATGGGCCCCATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG	1758
:			
:			
41	Db	GlyArgGlnThrAspGlyProIySerAspIleGluMetMetCysValMetSerThrGlu	60
:			
1759	Qy	GAAGCAGAGTTACCCATGAATGACCAACCGGTGAGTGGAGGTGGAAAGTGAATTTTGAT	1818
:			
61	Db	GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp	80
:			
1819	Qy	AGCGACAGAGATCTTACTAGAGATTATGCATCTCAGGTGGAATCAGATTCGGCGCTTACACAT	1878
:			
81	Db	SerGluGluIleuLeuAspIyAlaSerGlnValGluSerAspTrpProLeuThrHis	100
:			
1879	Qy	GGTCAATTTTTTCTCTATTTTCCGCAATTATGATTTCAGGTGGATACTTAGAGAAAGTGTAT	1938
:			
101	Db	GlyGlnPhePheSerIleLeuProIleIyAspSerGlyGlyIyLeuGluLysValIyTr	120
:			
1939	Qy	CAAACTCTCTAAATCGGTAGAGGCCAAACGGTTCACGATCGGATTTGTGCGCTTATCGTA	1998
:			
121	Db	GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal	140
:			
1999	Qy	GAAGAGCTGTTTGAATATGCAGGCCAAATGCGTAATATTCGTGTGCAAGGACCGACAACA	2058
:			
141	Db	GluGluLeuPheGluTrpAlaGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy	160
:			
2059	Qy	TTTCTACCATCTTGTACTGTACAGGTAGCAATGCGAGGTGCCATGTTTGATTGGTCTGCAT	2118
:			
161	Db	PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis	180
:			
2119	Qy	CATCGCAATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGAT	2178
:			
181	Db	HisArgIleCysIyThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp	200
:			
2179	Qy	CTTCCTTCAGGTTATGACCATCTGTGCCAGTTCGTATGTCGTGTCACCTTCGCACTCT	2238
:			
201	Db	LeuProSerGlyIyAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer	220
:			
2239	Qy	GAGAAACTTCTGGAATCGCTAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA	2298
:			
221	Db	GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg	240
:			
2299	Qy	CACGGATATATAGTGGATGTGTCAAAACCGATACCAATTT	2337
:			
241	Db	HisGlyIyIleValAspValSerLysArgIleProPhe	253

RESULT 13

```

PCT-US97-07853-4
; Sequence 4, Application PC/TUS9707853
; GENERAL INFORMATION:
; APPLICANT: Snyder, Linda A.
; APPLICANT: Satishchandran, C.
; TITLE OF INVENTION: CHIMERIC KANAMYCIN RESISTANCE GENE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/07853
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/642,045
; FILING DATE: 06-MAY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: APOL-0273
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US97-07853-4

Alignment Scores:
Pred. No.: 3 53e-119 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 1 Gaps: 0

```

US-10-681-086-1 (1-3156) x PCT-US97-07853-4 (1-253)

1579	QY	GTGAATGGACCAATATAATAGTACAGAGAAGAAAGATGTTCTTCATGAAATTAAG	1638
		:::	
1	Db	MetAsnGlyProIleileMetThrArgGluGluArgMetLysIleValHisGluIleLys	20
1639	QY	GAACGAATATTCGATAAATATCGGATGATGTAAAGGCTATTGGTGTTTTATGGCTCTCTT	1698
		:::	
21	Db	GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu	40
1699	QY	GGTCGTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG	1758
		:::	
41	Db	GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu	60
1759	QY	GAAGCAGAGTTCAGCCCATGAATGCACAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGAT	1818
		:::	
61	Db	GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysIleValAsnPheAsp	80
1819	QY	ACGGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAAATCAGATTGGCCGCTTACACAT	1878
		:::	
81	Db	SerGluGluIleLeuLeuAspTyrzrAsaSerGlnValGluSerAspTrpProLeuThrHis	100
1879	QY	GGTCAATTTTTTCTATTTTTCCGATTTTATGATTTCAGGTGGATCTACTAGAGAAGTGTAT	1938

Db 101 GlyGlnPheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAAATCTGTAATCGGTAGAGCCCAACGTTCCACGATCGATTTGTCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTTGAATATGAGGCAATGCGTAATATTCGTGTGCAAGGACGACAAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTCTACCATCTTGACTACAGGTAGCAATGCGCAGTGCATGTTGATGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGCATCTGTATACGACGAGCGCTTCGTCTTAACTGAAGCAGTAAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
Qy 2179 CTTCTCTTCAGGTTATGACCATCTGTGCGCAGTTCTGTAATGTCTGTCAACTTTCCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTCTCGAATCGCTAGAGAAATTCGGAATCGGATTCAGGATTCAGGACGACGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTyrAsnGlyIleGlnGlnTyrThrGluArg 240
Qy 2299 CACGGATATAGTGGATGTGCACAAACGATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 14

US-08-642-045A-4
; Sequence 4: Application US/08642045A
; GENERAL INFORMATION:
; APPLICANT: Snyder, Linda A.
; APPLICANT: Satishchandra, C.
; TITLE OF INVENTION: CHIMERIC KANAMYCIN RESISTANCE GENE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.0/6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,045A
; FILING DATE: 06-MAY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: APOL-0262
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-045A-4
Alignment Scores: 3.53e-119 Length: 253
Pred. No.: 1333.00 Matches: 252
Score:

Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 16 Gaps: 0
US-10-681-086-1 (1-3156) x US-08-642-045A-4 (1-253)
Qy 1579 GTCAATGGACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTCAATGAAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGAAATATTGATGATAAATATGGGATGATTTAAGGCTATTGGTGTATTGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATCGGCCTATTTCGGATATTGAGATGATGTGTGCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
Qy 1759 GAACGAGAGTTTCAGCCATGAATCGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAT 1818
Db 61 GluAlaGluPheSerHisGluTyrThrThrGlyLysValGluValAsnPheAsp 80
Qy 1819 AGCAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuAspTyrAlaSerGlnValGluSerAspTyrProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTGTCGATTTATGATTTCAGGTGCTGATCTTAGAGAAAGTGTAT 1938
Db 101 GlyGlnPheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
Qy 1939 CAAACTGCTAAATCGGTAGAGCCCAACCGTTCCACGATCGCATTTGTGCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTGCAATATGACGCAAAATGCGCTAATATTCTGTGTGCAAGGACGACGAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTCTACCATCTTGACTGACAGGTAGCAATGCGCAGTGCCTGCTGTTGTTGCTCTCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATCGCATCTGTATACGACGAGCGCTTCGTCTTAACTGAAGCAGTAAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
Qy 2179 CTTCTCTTCAGGTTATGACCATCTGTGCGCAGTTCTGTAATGTCTGCTCACTTTCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTCTGCAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGATTCAGGACGACGAACA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTyrAsnGlyIleGlnGlnTyrThrGluArg 240
Qy 2299 CACGGATATAGTGGATGTGCACAAACGATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
RESULT 15
US-09-791-537-104530
; Sequence 104530, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 104530
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-791-537-104530

Alignment Scores:
Pred. No.: 3.53e-119 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 27 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-791-537-104530 (1-253)

```
QY 1579 GTGAATGACCAATAATATGACTAGACAGAAAGAAATGAAGATTGTTTCATGAATAATGAAG 1638
Db :|||||
1 MetAenGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20

QY 1639 GAACGATATTGGATAATATGGGATGATGTTAAGGCTATTGGTTTATGGCTCTCTT 1698
Db :|||||
21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
Db :|||||
41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60

QY 1759 GAACGAGAGTTTCAGCCATGAATGGACAAACCGGTGAGTGGAGGTGGAAAGTGAATTTTGAT 1818
Db :|||||
61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp 80

QY 1819 AGCGAAGAGATTTCTACTAGATTATCATCTCAGGTGGATCAGATTGGCGCTTACACAT 1878
Db :|||||
81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100

QY 1879 GGTCAATTTTCTCTATTTCGCGATTATGATTTCAGGTGGATCTTAGAGAAAGTGTAT 1938
Db :|||||
101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120

QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTA 1998
Db :|||||
121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140

QY 1999 GAAGAGCTTTTGTATATGCAAGCAATGGCGTATATTCGTGTCAGAGGCCGACACACA 2058
Db :|||||
141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

QY 2059 TTTCACCATCTTGACTGTACAGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCAT 2118
Db :|||||
161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

QY 2119 CATCCATCTGTTATACGACGAGCCCTTCGGTCTTTAACTGAAGCAGTTTAAGCAATCAGAT 2178
Db :|||||
181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200

QY 2179 CTTCTCTCAGGTATGACCATCTGGCCAGTTCGTAAATGCTGTGTCAACTTTCGACTCT 2238
Db :|||||
201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

QY 2239 GAGAAACTTCTGGATCCGCTAGAGAAATTTCTGGATGGGATTCAGAGGTGGACAGAACGA 2298
Db :|||||
221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240

QY 2299 CACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db :|||||
241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

Search completed: February 10, 2006, 03:48:56
Job time : 1376 secs

QY 82 CGTAACTGGCGTCAATGGATGGAGCGCCGGTCCAGAGAGGAAATATGATGGCGAATAI 14

```
Db      21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
Qy      142 CAAACGGCTGCTGCTCAAAACAATATTATAGCGCTCGCAAGCATAGAGCTTTGATCGAT 201
Db      41 GlnThrValTrpSerSerAsnAsnTrpLeuGlyLeuAlaSerAspArgGluIleAsp 60
Qy      202 CGAGCCCAACAGCATTTGAGCAATTTGGGACAGGAAGCAGCGGTTTCAGTTTAAACGACA 261
Db      61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
Qy      262 GCGAATTCGGCTGGCATGAAAGCTAGAAAGAGATTGCGAGCTTTAAACTGACAGAA 321
Db      81 GlyAsnSerValTrpHisGlyLeuGluLysIleAlaSerPheLysLeuThrGlu 100
Qy      322 GGGGCGCTGCTGCTTTTCAGCGGTTACTTTGGCAATGTCGGTGCTCTTTTCATCCTTGCCA 381
Db      101 AlaAlaLeuLeuPheSerSerGlyTrpLeuAlaAsnValGlyValLeuSerSerLeuPro 120
Qy      382 GAAAGAGAGATGTCATTTTAAGTGACCGAGCTCAATCATGCAAGTATGATCGACGGCTGC 441
Db      121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
Qy      442 CGACTTCTTAGCGCTGATACAGTGTGTTATCGGCATATTGATATGATGAATGATCTTGAAGAA 501
Db      141 ArgLeuSerLysAlaAspThrValValTrpArgHisIleAspMetAsnAspLeuGluAsn 160
Qy      502 AAGCTGAATGAACACAGCGTTATCAGCGCGGTTTATCGTAACAGACGGAGTATTCAGC 561
Db      161 LysLeuAsnGlnThrGlnArgTrpGlnArgPheIleValThrAspGlyValPheSer 180
Qy      562 ATGGATGGCACAAATCGCCCTCTTGATCAGATCATCTCACTTTCGAAACGGCTATCATGCC 621
Db      181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTrpHisAla 200
Qy      622 TTCGTGCTGCTGATGATGCCACCAACAGAGATTTGGCGATTCGGGACAGGAACG 681
Db      201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
Qy      682 AGTGAATCTTTGGTGTTTGTCCGACATTTGTTATCGGCACCTTAAGCAAGCTGTGGC 741
Db      221 SerGluTrpPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
Qy      742 CGGAGAGAGGTTTTCGGCGAGGATCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db      241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
Qy      802 AGAATATTATCTTTCAAACCGCTATTCCGCCAGCAGCTGTGGCGCTGCTCAGAGGCT 861
Db      261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaHisGluAla 280
Qy      862 TTCACATCATTTGAAGCCAGCAGGGAAGAAACGACAGCTTTTATTTCTTATATCAGCATG 921
Db      281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
Qy      922 ATCAGAACCATGCTCAGGAATATGGGTATGTGTGAAAGAGAGATCACACACCATGATT 981
Db      301 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 320
Qy      982 CCTGTAGTCATTTGGCATGCCCATAAACCGTCTCTATTGCTGAAAACTCAGGGCAAG 1041
Db      321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
Qy      1042 GGAATTTATGCTCTGCTGCTGCGCCGCAACCGTTGCGCGGCTGAAGCCGAGTTTGA 1101
Db      341 GlyIleTyrAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 360
```

RESULT 2

US-60-752-355-47667

; Sequence 47667, Application US/60752355

; GENERAL INFORMATION:

; APPLICANT: Abad, Mark et al.

; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

; FILE REFERENCE: 38-21 (53720)B

; CURRENT APPLICATION NUMBER: US/60/752,355

; CURRENT FILING DATE: 2005-12-21

; NUMBER OF SEQ ID NOS: 52803

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 47667

; LENGTH: 371

; TYPE: PRN

; ORGANISM: Bacillus subtilis

US-60-752-355-47667

Alignment Scores:
Pred. No.: 7,04e-157 Length: 371
Score: 1676.00 Matches: 330
Percent Similarity: 93.3% Conservative: 6
Best Local Similarity: 91.7% Mismatches: 6
Query Match: 30.4% Indels: 18
DB: 8 Gaps: 1

US-10-681-086-1 (1-3156) x US-60-752-355-47667 (1-371)

```
Qy      22 TTGAGATTGATTCTCTGTTAAACGACGGTTAGACAGAATGAAGAAGCGCGGTACAT 81
Db      1 MetLysIleAspSerTrpLeuAsnAspArgLysAspIleAlaLysGluAlaGlyValHis 20
Qy      82 CGTAACCTCGCGTCAATGGATGGAGCCGCTTCCAGAGGAGAAATATTGATGGCGAAAT 141
Db      21 ArgAsnLeuArgSerMetAsnGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
Qy      142 CAAACGGTCTGGTCTCTCAAAACAATTTATTAGGGCTCCAGCGCATAGACGTTTGCATCGAT 201
Db      41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgGluIleAsp 60
Qy      202 GCAGCCCAACACAGCATTTGACGCAATTTGGACAGGAGCAGCGGTTACGCTTTAACGACA 261
Db      61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
Qy      262 GGCATTTCTGGCTCTGGCATGAAAGCTAGAAAGAGATTGCCAGCTTTAAACTCACAGAA 321
Db      81 GlyAsnSerValTrpHisGlyLysGluLysLeuGluLysIleAlaSerPheLysArgThrGlu 100
Qy      322 GCGGCGCTGCTGTTTTCGACGGTACTTTGCGCAATGTCGGTGCTCTTTCATCTGCCA 381
Db      101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
Qy      382 GAAAGAGAGATGTCATTTTAACTGACCGACTCAATCATGCAAGTATGATCGACGGCTGC 441
Db      121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerIleIleAspGlyCys 140
Qy      442 CGACTTCTTAAGGCTGATACAGTTGTTTATCGGCATATTGATGATGAATGATCTTTGAAAC 501
Db      141 ArgLeuSerLysAlaAspThrValValTrpArgHisIleAspMetAsnAspLeuGluAsn 160
Qy      502 AAGCTGAATGAAACACAGCGTTATCAGCGCGGTTTATCGTAACAGACGGAGTATTCAGC 561
Db      161 LysLeuAsnGlnThrGlnArgTrpGlnArgPheIleValThrAspGlyValPheSer 180
Qy      562 ATGATGGACCAATCGCCCTCTTGATCAGATCATCTCACTTTCGCAACCATCATGCC 621
Db      181 MetAspGlyThrIleValProLeuAspGlnIleSerLeuAlaLysArgTrpHisAla 200
Qy      622 TTCGTGCTGCTGATGATGCCACCAACAGGAGTTTTGGCGATTCGGGACAGGAACG 681
Db      201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyArgGlyThr 220
Qy      682 AGTGAATCTTTGGTGTGTTGTCGCCGACATTTGATCGAGCTTTCATCGACATTTTGTGGC 741
Db      221 SerGluTrpPheGlyValTyrProAspIleValIleGlyThrLeuSerLysAlaValGly 240
Qy      742 GCGAGAGAGGTTTTCGGCGAGGATCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db      241 ThrGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
Qy      802 AGAACATTTATCTTTCAAACCGCTATTCCGCCAGCAGCTGTGGCGCTGCTCAGAGGCT 861
Db      261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaHisGluAla 280
Qy      862 AGAACATTTATCTTTCAAACCGCTATTTCGCCAGCAGCTGTGGCGCTGCTCAGAGGCT 861
```



```
Db      8 LeuLysAspIleLysAsnLysGlyLeuTyrArgGluPheArgAsnValAsnAlaAlaGln 27
Qy      112 GTTCAGAGAGGAATATTGATGGGAAATCAAAACGGTCTGGTCTCAACAAATATTATTA 171
Db      28 GlyProTyrThrValIleAspGlyArgLysMetLeuMetSerAsnAsnTyrLeu 47
Qy      172 GGGCTCGCAAGCGATAGAGTTTATCATCGATCAGCCCAACAGCATTCAGCAATTTGGG 231
Db      48 GlyLeuCysAspAspIleArgLeuLysArgAlaIleGluSerIleArgLysPheGly 67
Qy      222 ACAGAAACGAGGTTTCAGTTTAAACAGACAGCAATTCGGTCTGGCATGAAGAGTAGAA 291
Db      68 ValGlyAlaGlyGlySerArgLeuThrCysGlyAsnPheGluLeuHieArgGluLeuGlu 87
Qy      292 AAGAAGATTGCCAGCTTTAAATCTGACAGAGAGCGGCTGCTGTTTCAGCGGTACTTG 351
Db      88 GluArgLeuAlaLysPheLysAspValGluSerCysIleValPheGlySerGlyTyrAla 107
Qy      352 GCCAATGTCGGTGTCTTTCATCTTCGACAGAAAGAAAGATGTCATTTAAAGTGACGAC 411
Db      108 AlaAsnIleGlyAlaIleSerGlyIleAlaAspLysAsnTrpValIlePheCysAspArg 127
Qy      412 CTCATCATGCAAGTATGATCGACGGCTGCCGACTTCTTAAGCGCTGATACAGTTGTTTAT 471
Db      128 LeuAsnHieAlaSerIleValAspGlyIleArgLeuSerGlyAlaLysLeuValValTyr 147
Qy      472 CGGCATATTGATGAATGATCTTGAAACCAAGCTGAATGAACACACGCTTATCAG--- 528
Db      148 LysHisCysAspMetGluAspLeuGluSerLysIle-----ValArgTyrHisThr 164
Qy      529 ---GCCCGCTTTATCGTAACAGACGGAGTATTACGATGATGGCACATCGCCCTCTT 585
Db      165 GlyLysSerLeuIleValThrAspGlyValPheSerMetAspGlyAspValAlaProVal 184
Qy      586 GATCAGATCATCTCATTCGAAACGCTATCATGCTTCGTTGCTGTTGATGATGCCAC 645
Db      185 AspArgIleValLysLeuAlaLysLysTyrAsnLeuMetThrMetValAspAspAlaHis 204
Qy      646 GCACAGAGATTGGGGCATTCGGGCAACAGGACGAGTGAATCTTTGGTGT--- 699
Db      205 AlaThrGlyIleLeuGlyGlyLysGlyArgGlyThrSerGluTyrPheGlyLeuLysAsp 224
Qy      700 TGTCCGACATGTTATCGCACCTTAAGCAAAAGCTGTTCGGCGGAGGAGTTTTCG 759
Db      225 AlaValAspIleSerMetGlyThrLeuSerLysAlaPheGlyValGluGlyGlyPheVal 244
Qy      760 GCAGATCAGCGGTCTTCATCGACTTTTGTGTAACCATGCCAGAACATTTATCTTTCAA 819
Db      245 AlaGlyLysArgLysLeuValAspPheLeuArgHisLysAlaLysSerPheIleTyrSer 264
Qy      820 ACCGCTATTCGCCAGCCAGCTGTGCGCTGCTCACAGGCTTTCACATCATTTGAAGCC 879
Db      265 ThrAlaProProHisAsnMetAlaAlaAlaLeuGluAlaLeuAsnIleIleGluThr 284
Qy      880 AGCAGGAAAAACGACAGCTTTATTTCTTATATATCATCATGATCAGAACCGCTGAAG 939
Db      285 GluProGlnAlaArgLysGluLeuAlaGluLysSerValTrpLeuArgAsnArgLeuIle 304
Qy      940 AATATGGTTATGTGGTGAAGAGAGATCACACCGATATTATTCCTGTAGTCAATTTGGCGAT 999
Db      305 GluLysGlyPheAsnValProLysGlyValThrProIleIleProLeuMetValGlyAsp 324
Qy      1000 GCCCATAAACGGTCTTATTTGCTGAAACAACTGACGGCAGGGAATTTATGCTCTCTGCC 1059
Db      325 ValAsnThrAlaValGluPheSerMetLeuLeuTyrAsnGluIleTyrIleProAla 344
Qy      1060 ATTGGCGCGCAACCGTTGCGCGGTTGAAGCGGATTCGA-----AGCTTCGGCGAGC 1113
Db      345 IleArgProProThrValProLysGlyThrSerArgLeuArgIleSerIleMetAlaSer 364
Qy      1114 AGGTCGAGATCAGGAATGAGTTT 1137
```

```
Db      365 HisSerTyrGluAspMetGluPhe 372
RESULT 5
US-60-752-355-26171
; Sequence 26171, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)/B
; CURRENT APPLICATION NUMBER: US/60752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26171
; LENGTH: 392
; TYPE: PRF
; ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-60-752-355-26171
Alignment Scores:
Pred. No.: 1,96e-74 Length: 392
Score: 843.00 Matches: 169
Percent Similarity: 66.8% Conservative: 77
Best Local Similarity: 45.9% Mismatches: 120
Query Match: 15.3% Indels: 2
DB: 8 Gaps: 1
```

US-10-681-086-1 (1-3156) x US-60-752-355-26171 (1-392)

```
Qy      37 TGGTTAAACGAGCGGTTAGACAGAAATGAAAGAACCGCGGTACATCGTAACCTCGGTCA 96
Db      7 PheLeuLysAlaGluLeuAsnSerMetLysGluAsnHisThrTrpGlnAspIleLysGln 26
Qy      97 ATGATGCGAGCGCGGTTCCAGAGAGGAATATTGATCGCGGAAATCAACGGTCTGGTCC 156
Db      27 LeuGluSerMetGlnGlyProSerValThrValAsnHisGlnLysValIleGlnLeuSer 46
Qy      157 TCAACAATATTATTAGGCGTCGCAAGGATAGAGCTTTGATCGATCGACGCCCAACAGCA 216
Db      47 SerAsnAsnTyrLeuGlyPheThrSerHisProArgLeuIleAsnAlaIleGlnLysAla 66
Qy      217 TTCAGCAATTTGGGACAGAACGCGGTTTCACGTTTAAACGACGACCAATTCGGTCTGG 276
Db      67 ValGlnGlnTyrGlyAlaGlyThrGlySerValArgThrIleAlaGlyThrPheThrMet 86
Qy      277 CATGAAAGCTAGAAAGAGATTGCCAGCTTTAAACTGACAGAGCGGCTCTGCTTT 336
Db      87 HieGlnGluLeuGluLysLysLeuAlaAlaPheLysLysThrGluAlaAlaLeuValPhe 106
Qy      337 TCAGACGGTTACTTGGCCAATGTCGGTGTCTCTTTTCATCTTCCAGAAAAGGAAGATGTC 396
Db      107 GluSerGlyPheThrThrAsnGlnGlyValLeuSerSerIleLeuSerLysGluAspIle 126
Qy      397 ATTTAAGTACGACGCTCAATCATGCAAGTATGATCGACGCGCTGCCGACTTTCTAAGGCT 456
Db      127 ValIleSerAspGluLeuAsnHisAlaSerIleIleAspGlyIleArgLeuThrLysAla 146
Qy      457 GATACAGTTGTTTATCGGCATATTGATGATGATCTTGAATAACAGCTGAATGAACA 516
Db      147 AspLysLysValTyrGlnHisValAsnMetSerAspLeuGluArgValLeuArgLysSer 166
Qy      517 CAGCGTTATCAGCGCGCTTTTATCGTAAACAGCGAGTATTTCAGCATGGATGGCACAACT 576
Db      167 MetAsnTyrArgMetArgLeuIleValThrAspGlyValPheSerMetAspGlyAsnIle 186
Qy      577 GCCCTCTTTCATCAGATCATCTCATTGCGAAACGCTATCATGCTTCGTTGCTGTTGAT 636
Db      187 AlaProLeuProAspIleValGluLeuAlaGluLysTyrAspAlaPheValMetValAsp 206
Qy      637 GATGCCCAACGACAGGAGTTTGGCGGATTCGGGACAGGAAACGAGTGAATACTTTGGT 696
Db      207 AspAlaHisAlaSerGlyValLeuGlyGluAsnGlyValArgGlyThrValAsnHisPheGly 226
```



```
QY 493 CTTGAAACAAAGCTG-----AATGAAACACAGCGTTATCAGCGCCGTTTATC 540
|||
Db 165 LeuArgAlaLysLeuAlaThrArgAlaAlaAspValAspGluAspGluGlnTrpLeuVal 184
|||
QY 541 GTAACACAGCGAGTATTCAGCATGGATGGACATATCCCTCTTGTATCAGATCATCTCA 600
|||
Db 185 ValThrAspSerValPheSerMetAspGlyAspIleAlaProLeuSerAlaIleCysAsp 204
|||
QY 601 CTTGCGAAACGCTATCATGCTCTGCTGCTGATGATCCACGCAACAGAGGTTTG 660
|||
Db 205 AlaValAspGluThrGlyAlaTrpLeuMetValAspGluAlaHisAlaThrGlyLeuPhe 224
|||
QY 661 GCGGATTCGGGACCAAGACGAGTGAATACTTTGGTGTGT-----CCCGACATTGTT 714
|||
Db 225 GlyAspSerGlyGlyValValGlnArgGluGlyLeuSerHisArgValAspIleGln 244
|||
QY 715 ATCGGACCTTAACCAAGCTGTGGCGGAGGAGGTTTGGCGGAGATCAGCGGTC 774
|||
Db 245 LeuGlyThrLeuSerLysAlaLeuAlaSerGlnGlyGlyTyrIleAlaGlyAspGluVal 264
|||
QY 775 TTCAATCAGCTTTTGTGTAACCATGCCAGAACATTTATCTTCAAAACCGCTATTCCGCCA 834
|||
Db 265 LeuIleGluThrLeuLeuSerAlaAlaArgSerPheValPheSerThrGlyLeuSerPro 284
|||
QY 835 GCGAGCTGTGCGGCTGCTCAGCGGCTTCAACATCATTTGAAGCCAGCAGGAAACGA 894
|||
Db 285 ProAsnAlaAlaAlaCysGluAlaLeuArgIleAlaArgLysThr---AspArgAla 303
|||
QY 895 CAGCTTTTATTTCTTATATCAGCATGATCAGAACCATCTGAAGAAATATGGTTATGTG 954
|||
Db 304 AlaGluLeuTrpAspThrValAlaThrLeuArgAspGlyLeuGluThrMetGlyTyrGlu 323
|||
QY 955 GTGAAGAGGATCACACACGATTTATCTGTAGTCAATTCGCGATGCGCATGCCATAAACGGTC 1014
|||
Db 324 ValLeuGlyGlu---ThrHisIleLeuProValIleValGlyAspArgGlyAspAlaLeu 342
|||
QY 1015 CTATTTGCTGAAACATGCGAGGCAAGGAATTTATCTCTGCTGCAATTCGCGCGCCAAAC 1074
|||
Db 343 GluLeuAlaAspArgLeuArgAspHisGlyIleValAlaProAlaIleArgProThr 362
|||
QY 1075 GTTGGCGCGGTTGAAGCCGGATTCGA 1101
|||
Db 363 ValProGluGlyThrSerArgIleArg 371
|||
RESULT 9
US-60-752-355-36490
; Sequence 36490, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36490
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Trichodesmium erythraeum IMS101
US-60-752-355-36490
Alignment Scores:
Pred. No.: 2,07e-62 Length: 393
Score: 721.50 Matches: 162
Percent Similarity: 62.4% Conservative: 85
Best Local Similarity: 40.9% Mismatches: 128
Query Match: 13.1% Indels: 21
DB: 8 Gaps: 7
US-10-681-086-1 (1-3156) x US-60-752-355-36490 (1-393)
QY 34 TCCTGGTTAAACGAGCGGTTAGACAGAAATGAAAGACCGCGGTACATCGTAACCTGCGG 93
```

```
Db 7 SerTrpIleGluSerSerLeuSerThrIleHisLeuAlaAsnTrpTyrArgSerValLys 26
|||
QY 94 TCAATCGATGAGCGCGGTTCCAGAGAGGAATATTGATCGCGAAATAACAAACGCTCGG 153
|||
Db 27 ThrIleGluSerIleProGlyAlaIleIleLysLeuGluGlyLysLysLeuIleAsnPhe 46
|||
QY 154 TCCTCAAAACAATTTAGGGCTCGCAAGCGATAGACGTTTGTATCGATGCGAGCCCAACA 213
|||
Db 47 AlaSerAsnAspTyrLeuGlyLeuAlaGlyAspGluArgLeuIleAlaIleGln 66
|||
QY 214 GCATTCACGAATTTGGGACAGAACGAGCGGTTTACGTTTAACTGACAGAGCGGCCTGCT 273
|||
Db 67 AlaThrLysGluPheGlySerGlySerThrGlySerArgLeuLeuThrGlyHisArgGlu 86
|||
QY 274 TGCATCAAAAAGCTAGAAAAGAGATTGCCAGCTTTTAACTGACAGAGCGGCCTGCT 333
|||
Db 87 IleHisGlnGlnLeuGluArgGluIleAlaLysLeuLysGlnThrGluSerAlaLeuVal 106
|||
QY 334 TTTTCAGCGGTTTACTTGGCCAATGTGCGTCTCTTTTCATCTTCCTGCGAGAAAGAGAT 393
|||
Db 107 PheSerSerGlyTyrLeuAlaAsnIleGlyValIleSerSerValValSerGlnArgAsp 126
|||
QY 394 GTCATTTTAACTACACGATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 453
|||
Db 127 LeuIleLeuSerAspGluTyrAsnHisSerSerLeuLysAsnGlyAlaIleLeuSerGly 146
|||
QY 454 GCTGATACAGTTCTTATCGGCATATTGATATGATATGATATGATATGATATGATATGATATG 513
|||
Db 147 AlaLysIleIleGluTyrSerHisAsnAsnIleGluTyrLeuLysAsnLysLeuGln 166
|||
QY 514 ---ACACAGCGTTATCAGCGCCGTTTATTCGTAAACACAGCGAGTATTCAGCATGGATGC 570
|||
Db 167 LysArgGluAsnTyrArgArgSerLeuIleIleThrAspSerValPheSerMetAspGly 186
|||
QY 571 ACATCGCCCTCTTATCGATCATCTCATCTCGGAAACGCTATCATGCTCTGCTGCTGCTC 630
|||
Db 187 AspLeuCysLysLeuProLeuLeuLeuLeuLeuLeuLeuLeuLysTyrAsnSerMetLeuLeu 206
|||
QY 631 GTTGATGATGCGCCACCAACAGGAGTTTGGGGGATTTCGGGCAACAGGAACAGAGTGAATAC 690
|||
Db 207 ValAspGluAlaHisAlaThrGlyValPheGlyIleAsnGlyGlyCysValGluHis 226
|||
QY 691 TTTGGTGTGTTGTCCTCCGAC-----ATTGTTATCGCACCTTAAAGCAAAGCTGTGGC 741
|||
Db 227 PheAsn---CysThrGlyLysGlnLeuIleGlnIleGlyThrLeuSerLysAlaLeuGly 245
|||
QY 742 GCGGAAGGAGGTTTTCGGGCGAGATCAGCGGCTTTCATCGACTTTTGTGTAACCATGCC 801
|||
Db 246 SerLeuGlyGlyTyrValAlaGlySerLysAsnLeuIleGluPheLeuArgAsnArgThr 265
|||
QY 802 AGAACATTTATCTTCAACCGCTATTTCGCCAGCCAGCTGTGCGGCTGCTCACAGGCT 861
|||
Db 266 ProThrTrpIleTyrThrThrGlyLeuThrProAlaAspThrAlaAlaLeuThrAla 285
|||
QY 862 TTCACATCATTTGAAGCCAGCAGGGAACACGACAGCTTTTATTTTCT----- 909
|||
Db 286 IleLysIleIleLysLysGluProGluArgArgMetLysLeuTrpGlnAsnLeuGluIle 305
|||
QY 910 TATATCAGCATGATCAGAAC-----AGTCTCAAGAAATATGGTTATGTTGTAAGGA 963
|||
Db 306 PheIleAsnLeuLeuGluThrGluSerGlnLeuHisLysGlyLysLysThrSerAsn 325
|||
QY 964 GATCACACCGGATTTCTCTGTAGTCATTGGCGATGCCCATAAACGGTCTATTGCT 1023
|||
Db 326 TyrGluSerProIleIleCysPheProLeuLysAsnAlaValGluAlaLeuLysValGly 345
|||
QY 1024 GAAAACTGCAAGGCAAGGAATTTATGCTCTGCCATTCGGCCGCCCAACGTTGCGCCG 1083
|||
Db 346 GluLysLeuLysGlnGluGlyIlePheAlaProAlaIleArgProThrValAsnThr 365
|||
QY 1084 GGTGAAAGCGGATTCGAAGCTTGGGACAGCAGCTCGAGATCAGGGAATGATTTATAAA 1143
|||
```

```
Db 366 SerArgileArgIleSerLeuMetSerThrHisGluThrSer----- 379
Qy 1144 TAAAAAAGCACCTGAAGAGTGCTTTTTTTTGTGATGGTTTGAACCTTG 1191
Db 380 -----HisLeuGlnGln---LeuileAlaAlaLeuileAsnLeu 391
RESULT 10
US-60-752-355-20828
; Sequence 20828, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20828
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Thermoplasma volcanium GSS1
US-60-752-355-20828
Alignment Scores:
Pred. No.: 1.8e-61 Length: 393
Score: 712.00 Matches: 153
Percent Similarity: 61.3% Conservative: 70
Best Local Similarity: 42.0% Mismatches: 137
Query Match: 12.9% Indels: 4
Gaps: 3
US-10-681-086-1 (1-3156) x US-60-752-355-20828 (1-393)
Qy 34 TCCTGGTTAAACGAGCGGTTAGACAGATGAAGAGCCGCGGTACATCGTAACCTGCGG 93
Db 5 AsnTrpValAspGluGluIleSerAlaLeuLysAlaGluGlyArgTyrValProileArg 24
Qy 94 TCAATGGATGAGCGCGGTTCCAGAGAGGAAATATGATGGCGAAATCAAAACGCTCGG 153
Db 25 ThrileGluSerAlaGlnGlyAlaTrpValLysIleAsnGlyLysGlnValLeuAsnMet 44
Qy 154 TCCTCAAAACAATTTATAGGCTCGCAAGCATAGACGCTTTGATCGATGCGAGCCCAACA 213
Db 45 CysSerAsnAnthyLysLeuGlyPheAlaAsnHisProGluThrLysLysAlaAlaIleGlu 64
Qy 214 GCATTTGACGAATTTGGGACAGAGAGCGGTTACCGTTTAAACGACGCAATTCGCTC 273
Db 65 AlaileGluGluTyrGlyValGlyAlaGlyAlaValArgSerIleAlaGlyThrAspGlu 84
Qy 274 TGGCATGAAGAGCTAGAAAGAGATTGCCAGCTTTAAACTGACAGAGCGGCTCTG 333
Db 85 IleHisIleLysLeuGluGluLysIleAlaSerPheLysHisSerGluAlaAlaLeuVal 104
Qy 334 TTTTCGAGCGGTACTTCGCCAATGTCGGTCTCTTTTCATCTCTGCCAGAAAGAGAGAT 393
Db 105 TyrGlnGlyGlyLeuLeuAlaAsnLeuGlyThrValProAlaLeuValGlyLysAspAsp 124
Qy 394 GTCATTTTAAATGACACGCTCAATCATCATCAAGTATGATCGAGCGGTCGCCACTTCTAAG 453
Db 125 IleilePheSerGluGluLeuAsnHisAlaSerIleileAspGlyThrArgLeuSerPro 144
Qy 454 GCTGNATCAGTTGTTATCGGCATATTGATGATGATGATGATGATGATGATGATGATGAT 513
Db 145 AlaLysArgIleValTyrLysHisLeuSerValAspAspLeuGluLysGlnAlaLysGlu 164
Qy 514 ACACAGCGTTAT---CACGCGGTTTATCGTAACAGAGAGGATATTTCAGCATGATGCG 570
Db 165 AsnArgSerSerGlyLysLeuAlaLeuValIleThrAspGlyValPheSerMetAspGly 184
Qy 571 ACAATCGCCCTCTTGTATCAGATCATCATCTTCGGAACCGCTATCATGCGCTTCGTGTC 630
Db 185 AspIleAlaProLeuProGluIleValAspIleAlaGluLysTyrAspIleMetThrTyr 204
```

```
Qy 631 GTTGATGATCCACGCAACAGGAGTTTGGCGGATTCGGGACAAGCAAGCAAGTGAATAC 690
Db 205 ValAspAspAlaHisGlyGluGlyValLeuGlyAspHisGlyArgGlyValLeuAsnTyr 224
Qy 691 TTTGGTGCTT-----TGTCCCGCAGCATTTATCCGACACCTTTAAGCAAAAGCTGTTGGCGCG 744
Db 225 PheHisLeuGluAspLysValAspIleGluMetGlyThrPheSerLysAlaLeuGlySer 244
Qy 745 GAAGGAGTTTTCGGGAGGATCAGCGGTCTTCATCGACTTTTGTCTGAACCATGCCAGA 804
Db 245 MetGlyGlyPheValAlaGlySerAlaGluMetIleAspLeuLeuLysGlnLysAlaArg 264
Qy 805 ACATTTATCTTCAACCGCTATTCCCGCAGCCAGCTGTGGCGTGTCTCAGGAGCTTTC 864
Db 265 ProPheLeuPheSerSerAlaLeuAsnProGlyAspAlaAlaValLeuLysAlaIle 284
Qy 865 AACATCATTTGAAGCCAGCAGGAAAAACAGACGCTTTTATTTTATATATATCAGCATGATC 924
Db 285 GluileLeuGluLysAspSerLeuValLysLysLeuTrpGluAsnAlaAspLeuLeu 304
Qy 925 AGAACCATGCTGAAGATATGGGTATATGGTGAAGAGATCACACACCATTTATTCCT 984
Db 305 LysLysSerLeuAlaAspLeuGlyTyrAsnThrGlyHisSerLysThrProileThrPro 324
Qy 985 GTAGTCATTTGGCGATGCCCATAAACGCTCTATTGCTGAAAAAAGCTG---CAGGCGCAAG 1041
Db 325 ValIleileGlyAspGluLysLysThrValGluLeuSerThrLysLeuTyrGluGluLys 344
Qy 1042 GGAATTTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1101
Db 345 AsnValPheAlaSerProileValTyrProThrValProLysGlyThrAlaArgIleArg 364
Qy 1102 AGCTTGGCGCAGC 1113
Db 365 LeuMetProSer 368
RESULT 11
US-60-752-355-13501
; Sequence 13501, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13501
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Thermoplasma volcanium GSS1
US-60-752-355-13501
Alignment Scores:
Pred. No.: 1.81e-61 Length: 396
Score: 712.00 Matches: 153
Percent Similarity: 61.3% Conservative: 70
Best Local Similarity: 42.0% Mismatches: 137
Query Match: 12.9% Indels: 4
Gaps: 3
US-10-681-086-1 (1-3156) x US-60-752-355-13501 (1-396)
Qy 34 TCCTGGTTAAACGAGCGGTTAGACAGATGAAGAGCCGCGGTACATCGTAACCTGCGG 93
Db 8 AsnTrpValAspGluGluIleSerAlaLeuLysAlaGluGlyArgTyrValProileArg 27
Qy 94 TCAATGGATGAGCGCGGTTCCAGAGAGGAAATATGATGGCGAAATCAAAAGCTGCTGG 153
Db 28 ThrileGluSerAlaGlnGlyAlaTrpValLysIleAsnGlyLysGlnValLeuAsnMet 47
Qy 154 TCCTCAAAACAATTTATGAGGCTCGCAAGCGATGATGATGATGATGATGATGATGATGAT 213
```

Db 48 CysSerAsnAsnTyrLeuGlyPheAlaAsnHisProGluThrLysLysAlaAlaIleGlu 67
QY 214 GCATTCACGCAATTTCCGACAGCAAGCAGCGGTTACGTTTAAACGACGCGCAATTCGGTC 273
Db 68 AlaIleGluGluTyrGlyValGlyAlaGlyAlaValArgSerIleAlaGlyThrAspGlu 87
QY 274 TGCATGAAAGCTAGAAAAGAGATTGCGAGCTTTTAAACTGACAGAGCGCCCTCGTG 333
Db 88 IleHisIleLeuGluGluGlyLysLysAlaSerPheLysHisSerGluAlaLeuVal 107
QY 334 TTTTCAGCGGTACTTGGCCATGTCGGTCTGCTTTCATCTTCCAGCAAGAGAGAT 393
Db 108 TyrGlnGlyLeuLeuAlaLeuLeuValProAlaLeuValGlyLysAspAsp 127
QY 394 GTCATTTTAAGTACGACGCTCAATCATGCAAGTATGATGACGCGTCCGCGACTTCTAAG 453
Db 128 IleIlePheSerGluGluLeuAsnHisAlaSerIleAspGlyThrArgLeuSerPro 147
QY 454 GCTGATACAGTTGTTTATCGGCATATTGATATGATCTTGAACAACAGCTGAATGAA 513
Db 148 AlaLysArgIleValTyrLysHisLeuSerValAspAspLeuGluLysGlnAlaLysGlu 167
QY 514 ACACAGCGTTAT---CAGCGCCGTTTATCGTAAACAGAGAGTATTCAGCATGCGC 570
Db 168 AsnArgSerSerGlyLysLysAlaLeuValIleThrAspGlyValPheSerMetAspGly 187
QY 571 ACATCGCCCTCTGATCAGATCACTCATCTGCGAAGCGCTATCATGCTTCGTTGGTC 630
Db 188 AspIleAlaProLeuProGluIleValAspIleAlaGluLysTyrAspIleMetThrTyr 207
QY 631 GTTGATGATGCCCAACAGCAGAGTTTGGCGGATTTCGGGCAACAGGAGCTGAATAC 690
Db 208 ValAspAspAlaHisGlyGluGlyValLeuGlyAspHisGlyArgGlyIleValAsnTyr 227
QY 691 TTTGGTGT---TGTCCGACATTGTTATCGGCACCTTAAGCAAAAGCTGTTCGCGC 744
Db 228 PheHisLeuGluAspLysValAspIleGluMetGlyThrPheSerLysAlaLeuGlySer 247
QY 745 GAAGGAGTTTTCGCGCAGGATCAGCGTCTTCATCGACTTTTGTGTAACCATGCCAGA 804
Db 248 MetGlyGlyPheValAlaGlySerAlaGluMetIleAspLeuLeuLysGlnLysAlaArg 267
QY 805 ACATTTATCTTTCAACCGCTATTTCGCGCAGCCAGCTGTGCGGCTGTCTACGAGGCTTC 864
Db 268 ProPheLeuPheSerSerAlaLeuAsnProGlyAspAlaAlaValLeuLysAlaIle 287
QY 865 AACATCATGAACGACGAGGAAAACAGACAGCTTTATTTTCTTATATCATCATGATC 924
Db 288 GluIleLeuGluLysAspSerLeuValLysLysLeuTrpGluAsnAlaAspLeuLeu 307
QY 925 AGAACGCTCTGAAGATATGGCTTATGCTGTAAGGAGATCACACACCGGATTTCTCT 984
Db 308 LysLysSerLeuAlaAspLeuGlyTyrAsnThrGlyHisSerLysThrProIleThrPro 327
QY 985 GTAGTCACTTCGCGATGCAAAACGCTCTATTGCTGAAACAACTG---CAGGCGCAAG 1041
Db 328 ValIleIleGlyAspGluLysLysThrValGluLeuSerThrLysLeuTyrGluLys 347
QY 1042 GGAATTTATCTCTGCGATTCGCGCCCAACCGTTGCGCGCGGTAACCGGATTCGA 1101
Db 348 AsnValPheAlaSerProIleValTyrProThrValProLysGlyThrAlaArgIleArg 367
QY 1102 AGCTTGGCGCAGC 1113
Db 368 LeuMetProSer 371

RESULT 12

US-60-752-355-4236
; Sequence 4236, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B

; CURRENT APPLICATION NUMBER: US/60/752,355;
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4236
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Thermus thermophilus HB27
US-60-752-355-4236

Alignment Scores:
Pred. No.: 12e-60 Length: 452
Score: 704.00 Matches: 155
Percent Similarity: 60.5% Conservative: 72
Best Local Similarity: 41.3% Mismatches: 142
Query Match: 12.8% Indels: 6
DB: 4 Gaps: 4

US-10-681-086-1 (1-3156) x US-60-752-355-4236 (1-452)

QY 1 GGATCCACGAGGTAA---CGAGCCTTGAAGATTGAT-----TCCTGGTTAAACGAGCGG 51
Db 50 GlyProArgArgLeuAspLysAlaMetSerLeuAspLeuArgAlaArgValArgAspGlu 69
QY 52 TTAGACAGAAATGAAGAAGCGCGGTACATCGTAACCTCGGTCATGATGATGAGAGCGCG 111
Db 70 LeuGluArgLeuLysArgGluGlyLeuTyrIleSerProLysValLeuGluAlaProGln 89
QY 112 GTTCCAGAGAGNATATTGATGCGGAAATCAACCGTCTGCTCTCAACCAATTTTAA 171
Db 90 GluProAlaThrArgValGluGlyArgGluValValAsnLeuAlaSerAsnAsnTyrLeu 109
QY 172 GGGCTCGCAAGCGATAGACGTTTGTATCGATCGAGCCCAACAGCATTCGACCAATTTGGG 231
Db 110 GlyPheAlaAsnHisProTyrLeuLysGluLysAlaArgGlnTyrLeuGluLysTrpGly 129
QY 232 ACAGGAAGCAGCGGTTTCAGCTTTAAACGACGAGCAATTCGCTCGGATGAAAGCTAGAA 291
Db 130 AlaGlySerGlyAlaValArgThrIleAlaGlyThrPheThrTyrHisValGluLeuGlu 149
QY 292 AAGAAGATTGCCAGCTTTAACTGACAGAGCGGCTGCTGTTTCGAGCGGTACTTG 351
Db 150 GluAlaLeuAlaArgPheLysGlyThrGluSerAlaLeuValLeuGlnSerGlyPheThr 169
QY 352 GCCAATGTCGCTGCTCTTCATCTTCCTGCAAGAAAGAGAGATGCTCATTTTAAGTGACCAG 411
Db 170 AlaAsnGlnGlyValLeuGlyAlaLeuLeuLysGluLysValPheSerAspGlu 189
QY 412 CTCATATCATGCAAGTATGATCGAGCGCTGCCGACTTTCTAAGCTGATACAGTTGTTAT 471
Db 190 LeuAsnHisAlaSerIleIleAspGlyLeuArgLeuThrLysAlaThrArgLeuValPhe 209
QY 472 CGGCATATTGATATGATGATCTTGAACAAACAGCTGAATGAACACAGCGTTATCAGCGC 531
Db 210 ArgHisAlaAspValAlaHisLeuGluLeuLysAlaHisAspThrAspGlyLeu 229
QY 532 CGTTTATCGTAAACAGCGAGTATTTCAGCATCGATGGCAACAATCGCCCTCTTGTATCAG 591
Db 230 LysLeuIleValThrAspGlyValPheSerMetAspGlyAspIleAlaProLeuAspLys 249
QY 592 ATCATCTCATCTGCGAAACGCTATCATGCTTCGCTGCTGCTGCTGATGATGCGCCAGCA 651
Db 250 IleValProLeuAlaLysLysTyrArgAlaValValTyrValAspAspAlaHisGlySer 269
QY 652 GGAGTTTTCGGCGATTTCGGGACAGGACGAGTGAATCTTTCGTT-----TGTCCC 705
Db 270 GlyValLeuGlyGluLysGlyLysGlyThrValHisPheGlyPheHisGlnAspPro 289
QY 706 GACATTGTT---ATCGGCACCTTAAAGCAAGCTGTTGGCGCGGAAAGAGGAGTTTTCGCGCA 762
Db 290 AspValIleGlnValAlaThrLeuSerLysAlaTrpAlaGlyIleGlyGlyTyrAlaAla 309
QY 763 GGATCAGCGGTCTTCATCGACTTTTTCGTAACCATGCCAGAACATTTATCTTTCAAACC 822

Db 322 ThrProileThrProValLeuPheGlyGluAlaProLeuAlaPheGluAlaSerArgLeu 341
 QY 1030 CTGACGGCAGGAAGAAATTATGTCTCTGCCATTGGCCGCCCAACCGTTGCGCCGGTGAA 108
 Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 361
 Db 342 LeuLeuGluGluGlyValPheAlaValGlyIleGlyPheProThrValProArgGlyLys 361
 QY 1090 AGCCGGATTGCAAGCTTG 1107
 Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 367
 Db 362 AlaArgIleArgAenIle 367
 RESULT 15
 US-60-752-355-35003
 ; Sequence 35003, Application US/60752355
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: US/60752,355
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 35003
 ; LENGTH: 393
 ; TYPE: PRT
 ; ORGANISM: Thermoplasma acidophilum DSM 1728
 US-60-752-355-35003
 Alignment Scores:
 Pred. No.: 2,78e-60 Length: 393
 Score: 700.00 Matches: 154
 Percent Similarity: 59.3% Conservative: 64
 Best Local Similarity: 42.3% Mismatches: 142
 Query Match: 12.7% Indels: 4
 DB: 8 Gaps: 3
 US-10-681-086-1 (1-3156) x US-60-752-355-35003 (1-393)
 QY 34 TCTGTGTTAAACGAGCGGTTAGACAGAAATGAAAGACCGCGGTACATCGTAACTCGCGG 93
 |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 24
 Db 5 SerTrpValGluGluLeuSerAlaLeuLysAlaGluGlyArgTyrValProIleArg 24
 QY 94 TCAATGGATGGAGCGCGGTTCCAGACAGGAATATGATGCGGAAATCAACCGCTCTGG 153
 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 44
 Db 25 ThrIleGluSerAlaGlnGlySerTrpValThrIleGlyLysLysValLeuAsnMet 44
 QY 154 TCTCTCAACAAATATTATTAGCGCTCGCAAGCGATAGACGTTTGATCGATCGACGCCCAACA 213
 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 64
 Db 45 CysSerAsnAsnTyrLeuGlyPheAlaAsnHisProGluThrLysLysAlaAlaIleGlu 64
 QY 214 GCATTGCGACCAATTGGGACAGGAAGCAGCGGTTTCACGTTTAAACGACAGCAATTCGGTC 273
 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 84
 Db 65 AlaIleGluGlnTyrGlyValGlyAlaGlyAlaValArgSerIleAlaGlyThrAspGlu 84
 QY 274 TGGCATGAAAGCTAGAAAAGAAATGTCACGCTTTAACTGACAGAGCGCGCTCGTG 333
 |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 104
 Db 85 IleHlaIleGluLeuGluGluLysIleAlaLysPheLysHisMetGluSerAlaLeuVal 104
 QY 334 TTTTCGAGCGGTACTTGGCCAATGTCGGGTGCTCTTCATCCTTGGCCAAAAGAAAGAT 393
 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 124
 Db 105 TyrGlnGlyGlyLeuLeuAlaAsnValGlyThrIleProAlaLeuValGlyLysAsp 124
 QY 394 GTCATTTTAAGTACCAGCTCAATCATCGCAAGTATGATCGACGCGTCCCGACTTTCTAAG 453
 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 144
 Db 125 ValIlePheSerGluGluLeuAsnHisAlaSerIleIleAspGlyThrArgLeuSerSer 144
 QY 454 GCTGATACAGTTGTTTATCGGCATATTGATATGATATGATCTTGTAAACAACTGAATCAA 513
 |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 164
 Db 145 AlalyeArgIleValTyrLysHisLeuSerValGluAspLeuGluLysGlnIleArgGlu 164
 QY 514 ACACAG--CGTTATACGCGCGTTTTTATCGTAAACAGACGAGTATTCAGCATGATCGCG 570
 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 184
 Db 165 AsnArgSerAlaPheLysLysAlaLeuValIleThrAspGlyValPheSerMetAspGly 184

[illegible]


```
Qy 571 ACAATCGCCCTCTTGATCAGATCATCTCACTTGGAAACGCTATCATGCTCTGCTGTC 630
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
185 AspilealaProLeuProGluIleThrGluValAlaGluLysAsnAspValMetValTyr 204
Qy 631 GTTGATGATGCCACGCAACAGGAGTTTGGCGGATTTCGGACACAGGAACGAGTGAATAC 690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
205 ValAspAspAlaHisGlyGluGlyValLeuGlyAspHisGlyArgGlyIleValAsnTyr 224
Qy 691 TTTGGTGT-----TGTCGCCGACATGTTATCGGCACCTTAAGCAAGCTGTGGCGC 744
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
225 PheHisLeuGluAspArgValAspIleGluMetGlyThrPheSerLysAlaLeuGlySer 244
Qy 745 GAAGGAGGTTTTCGGCGGAGGATCAGCGCTCTTCATCGACTTTTGTCTGAACCATGCCAGA 804
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
245 MetGlyGlyPheValAlaGlySerAlaAspLeuIleAspLeuLysGlnLysAlaIle 264
Qy 805 ACATTTATCTTTCAAAACCGCTATTCCGCCAGCCAGCTGTGCGGCTGCTCACGAGCTTTC 864
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 ProPheLeuPheSerSerAlaLeuAsnProGlyAspAlaAlaValLeuLysAlaIle 284
Qy 865 AACATCATTTGAAGCCAGCAGGAAACGACAGCTTTTATTTCTTATATCAGCATGATC 924
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
285 GluIleLeuGluLysAspSerLeuIleLysLysLeuTrpHisAsnSerAspIleLeu 304
Qy 925 AGAACCACTCTGAAGAATATGGTTATGTGTGTAAGGAGATCACACACCGATTATTCT 984
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
305 LysLysSerLeuSerAspValGlyTyrAsnThrGlyHisSerLysThrProIleThrPro 324
Qy 985 GTAGTCATTGCGGATGCCCATAAACCGTCTCTATTGCTGAAAACTG---CAGGGCAAG 1041
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
325 ValIleIleGlyAspGluLysLysThrValGluLeuSerLysMetLeuTyrAspGluLys 344
Qy 1042 GGAATTTATGCTCTGCCATTTCGGCCGCCAACCGCTTGGCCGGGTGAAGCCGGATTGCA 1101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
345 AspValPheAlaSerProIleValTyrProThrValProLysGlyThrAlaArgIleArg 364
Qy 1102 AGCTTGGGCAGC 1113
Db |||||
365 LeuMetProSer 368
```

Search completed: February 10, 2006, 03:51:30
Job time : 93 secs